L 37.8: P69/2

A Guide for Planning UI Program Improvement Studies

U.S. Department of Labor
Employment and Training Administration
June 1992

VL BEST COPY AVAILABLE COMPLETED

BLANK PAGE

Material contained in this publication is in the public domain and may be reproduced, fully or partially, without permission of the Federal Government. Source credit is requested but not required. Permission is required only to reproduce any copyrighted material contained herein.

This material will be made available to sensory impaired individuals upon request.
Voice phone: 202-535-0604
TDD* phone: 1-800-326-2577

*Telecommunications Device for the Deaf.

A

A Guide for Planning UI Program Improvement Studies

U.S. Department of Labor
Lynn Martin, Secretary

Employment and Training Administration
Robert T. Jones
Assistant Secretary of Labor

Unemployment Insurance Service
Mary Ann Wyrsch, Director
1992

This technical assistance guide (TAG) was prepared for the Unemployment Insurance Service, U.S. Department of Labor, by staff of the Bureau of Unemployment Compensation, Pennsylvania Department of Labor and Industry. It is designed to be a practical guide for use by State employment security agencies (SESAs) in the design and implementation of Unemployment Insurance program improvement studies.

The TAG is a major product of a special project funded by the Department in September 1990 to provide technical assistance to SESAs on the planning and conduct of short-term UI program studies. Pennsylvania's Bureau of Unemployment Compensation administered this project from October 1990 through March 1992.

UIS

A GUIDE FOR PLANNING UI PROGRAM IMPROVEMENT STUDIES

Prepared by:

Pennsylvania Department of Labor and Industry
Bureau of Unemployment Compensation
Richard A. Puerzer, Director
Harrisburg, PA 17210

Grantee authors:

Robert J. Dougherty, principal author, assisted by Bureau colleagues Stephen L. Hinnenkamp, and Bryan M. Diehl.

Submitted to:

U.S. Department of Labor
Employment and Training Administration
Unemployment Insurance Service
200 Constitution Ave., N.W.
Washington, D.C.

ACKNOWLEDGMENTS

Project staff wish to acknowledge the contributions of the following individuals for their able and valuable assistance:

> Contract consultants Paul L. Burgess, PH.D. and Robert D. St. Louis, Ph.D. of Burgess and Associates, Scottsdale, Arizona.

In addition, we thank Federal staff in the Philadelphia Regional Office and the National Office of the Employment and Training Administration, U.S. Department of Labor, for their helpful guidance and support in the implementation of the Unemployment Insurance Quality Control Program Improvement Technical Assistance Project.

The project team assumes responsibility for any omission or inaccuracy in this document.

Robert J. Dougherty
Project Administrator

PREFACE

Numerous technical books have been written on the issues involved in conducting program improvement (PI) studies. These books deal with all of the technical issues addressed in this Technical Assistance Guide (TAG) including experimental design, data collection, data analysis, benefit/cost analysis, sample survey design, exploratory data analysis and the other issues that arise in completing program studies. The goal of this TAG is not to provide another such book. Instead, issues that can keep a PI study from achieving its objectives are emphasized. Thus, how various issues can affect the successful completion of a study are explained. However, there is no attempt to explain these issues in the kind of depth devoted to these topics in other books that are widely available. For example, the importance of considering various experimental design issues that relate to PI study development is explained, but the complex details of actually selecting a particular experimental design are not explained. When explaining why an issue is important, however, reference sources that elaborate on the issue are provided for the interested reader. These sources include very detailed analyses of the issues covered in this TAG.

The TAG is organized around a series of checklists that identify key issues involved in planning and conducting UI program improvement studies. Each chapter begins with a checklist. The remainder of the chapter translates the checklist into a UI context, and explains why the items on the checklist are important for PI studies.

The major goal of this TAG is to provide UI personnel with a working knowledge of the issues that must be considered to design and complete useful PI studies. To assist in achieving this goal, each step in the PI study process is related to a concrete UI study example that is referenced throughout the TAG. The example used is explained in Chapter II. It is designed to illustrate most of the technical issues that are likely to be involved in most PI studies, not just the type of study discussed in the example. At the outset, it should be emphasized that the PI study development process sometimes can be completed easily, without dwelling on some of the topics contained in the TAG. Nonetheless, careful review of the chapter checklists should be useful in ensuring that all steps usually necessary for an effective study are considered.

Chapter I provides some background on the Benefits Quality Control (BQC) program and the program improvement process. Then the UI study example used to illustrate the PI process throughout this TAG is developed in Chapter II. It assumes that the UI agency wants to address the problem of payment errors caused by incorrect partial payments. Chapter II shows how UI data sources can be used to identify the magnitude of this problem. Chapter III illustrates how agency personnel might work together to develop realistic study objectives. Chapters IV through IX illustrate how an agency might prepare and implement a PI study proposal. Chapter X discusses some issues involved in taking action based on the results of the PI study.

The TAG was prepared in this "linear" format to present the material in an organized fashion. But planning a study actually is a more "circular" process. Applying the concepts in this TAG to an actual study obviously is more complicated and less structured than it may appear from the discussion in the TAG.

An advantage of integrating a single example throughout the TAG is that all phases of the PI study process can be illustrated.

The reader is thus given a complete picture of the PI study process. An advantage of the particular example selected (incorrect partial payments) is that it can be used to illustrate how to conduct most of the types of PI studies States may wish to consider. Although the data and situations analyzed in the TAG are hypothetical, actual UI program findings were consulted to add some realism. But the example was chosen primarily to be illustrative, not to provide a practical guide for a partial earnings PI study for any particular SESA. Also, the example necessarily is incomplete in terms of study scope, planning and implementation. Nonetheless, it should be useful in illustrating the various aspects of the PI study process.

BLANK PAGE

CONTENTS

Page

Preface iii

Chapter I. Program Improvement in the UI System

A Brief History of Quality Control I-1
Background of UI Quality Control I-2
SESA Special Studies I-4
Emergence of QC Program Improvement Studies. I-5
Guideline for QC-Supported PI Studies. I-7
Use of the Technical Assistance Guide (TAG) I-8

Chapter II. Problem Identification

Checklist II-1
Introduction II-3
UI Study Example II-4
Utilizing Existing UI Data II-6
Quality Control Data II-6
Secondary Data Sources II-9
Data Source Matrix II-10
Limitations of Nonrandom Samples II-12
Building Organizational Support. II-13
Problem Cause and Impacts II-16
Estimating Costs II-16
Assessing Solvability; Potential Barriers. II-18
Problem Identification Report. II-19

Chapter III. Study Goals and Objectives

Checklist III-1
Study Purpose. III-2
System Boundaries. III-3
Study Objectives III-5
Study Feasibility. III-7

Chapter IV. Study Approach

Checklist IV-1
Validity Threats IV-4
Types of Studies IV-7
Selection of Feasible Approaches IV-10
Checks for Biases and Confounding. IV-15
Generalizability IV-16
Selection of the Study Approach. IV-17

Chapter V. Study Preparation

Checklist V-1
What Types of Data to Collect. V-3
How Much Data to Collect V-8
How Should the Data Be Collected V-11
Schedule of Activities V-14
Study Management and Control V-16
Threats to a PI Study. V-17

Chapter VI. Proposal Preparation

Checklist VI-1
Problem Identification VI-3
Study Objectives and Methodology VI-4
Administrative Considerations VI-4
Minimum PI Proposal Requirements (Table) VI-6

Chapter VII. Data Collection and Recording

Checklist VII-1
Design Accuracy VII-3
Collection Accuracy VII-4
Encoding Accuracy VII-4

Chapter VIII. Data Analysis

Checklist VIII-1
Sampled vs. Target Population VIII-3
Reliability and Validity VIII-8
Presentation of Results VIII-11
Recommended Actions VIII-12

Chapter IX. Final Report

Checklist IX-1
Cover Page IX-2
Executive Summary IX-2
Background/Problem Assessment IX-3
Description of Study Approach and Data Collection . . IX-3
Analysis and Findings IX-3
Conclusions and Recommendations IX-4
Illustrative Final Study Report Checklist (Table). . . IX-6

Chapter X. Implementation and Follow-up

Checklist X-1
Implementation Plan X-3
Potential Barriers. X-3
Identifying Barriers to Program Improvement Action. . X-4
Assessing the Impact of the Action X-5

Appendix A

UIPL 37-91. . . . . A-1
Bibliography. A-8
Glossary. A-10

Appendix B

Sample Letter to Regional Office Regarding PI Study Proposal (BYE). B-1
Sample PI Study Proposal. B-2
Sample Cover Letter to Regional Office (BRI). B-16
Sample PI Study Proposal (Abbreviated). B-17

Appendix C

Case Study 1. . . . . C-1
Case Study 2. . . . . C-3
Case Study 3. . . . . C-5
Case Study 4. . . . . C-7
Case Study 5. . . . . C-9
Case Study 6. . . . . C-10
Design Flowchart. C-12
Sample Size Chart C-13

BLANK PAGE

LIST OF TABLES AND CHARTS

Table/Chart Page

Table II-1 - Problem Identification Data Source Matrix . . II-21
Table II-2 - Organization Coordination Matrix II-22
Table II-3 - Organizational Support Table II-23

Chart III-1 - Study Variable Diagram. III-8

Table V-1 - Summary of Study Variables for BYE Example . . V-20
Table V-2 - Alternative Sample Sizes For the BYE Reporting Procedures Study V-21
Table V-3 - Advantages and Disadvantages of Alternative Data Collection Methods V-22
Table V-4 - Guide for Questionnaire Construction V-23
Table V-5 - Some Major Activities for Example PI Study . . V-25
Table V-6 - Major Resources Available for Example PI Study V-26
Table V-7 - Estimate Resource Requirements for Major Activities for Example PI Study V-27

Chart V-1 - Gantt Chart for Major Activities for Example PI Study V-28

Table VI-1 - Minimum Requirements for PI Proposal VI-6

Table VIII-1 - 95% Confidence Interval for the Effect of the BYE Test Procedure VIII-17
Table VIII-2 - Some Cost-Benefit Considerations VIII-18
Table VIII-3 - CBA Worksheet VIII-20

Figure VIII-1 - Confidence Interval for BYE Example With and Without Using the Linear Model . VIII-21

Table IX-1 - Illustrative Final Report Outline IX-6

xi

BLANK PAGE

CHAPTER I

PROGRAM IMPROVEMENT IN THE UI SYSTEM

1. **A Brief History of Quality Control**. A major goal established for the Unemployment Insurance (UI) system during the early 1980s was to develop a reliable, uniform mechanism for assessing the accuracy of benefit payments, State by State. The program mechanism that emerged in 1986 for this purpose is known as Benefits Quality Control (BQC).

The concept and assumptions of quality control for any work process were formulated over 80 years ago. However, W.E. Deming, an American statistician, is given credit widely for bringing the concept to prominence in commerce and industry in the late 1940s.

Realizing a need to improve the quality of manufactured products, Deming promoted a system of continuously drawing statistically selected samples of finished products to determine whether the whole batch was meeting production standards. If too few of the sampled products met established standards, the production process required adjustments. Further samples were then drawn to determine whether the adjustments made brought production to an acceptable level of performance. Generally, quality control (QC) involves the sampling of products or services, investigating their caliber or effectiveness, and correcting any process or judgment affecting their quality. The emphasis of QC is placed on the prevention of errors rather than the inspection of the total output of an operation for flawed products or ineffective services.

The Federal government adopted the QC process with its application to public family welfare and food stamp programs over 20 years ago. While the resulting designs for QC have differed widely depending upon the nature of each program, the basic principles are fairly well agreed upon.

2. **Background of UI Quality Control**. The basic QC concept of continuous intensive investigation of a small but representative sample of cases was first applied to UI in the late 1970s. Under the aegis of the National Commission of Unemployment Compensation, an experimental system for verifying payments was tested in six cities. This experiment determined that the actual error rate among UI payments in the test sites was several times what the then current UI system of error rate measurement indicated. It demonstrated that in order to estimate the UI payment error rate accurately, field verifications had to supplement record reviews.

Intrigued by the findings of this experiment, the Department modified and systematized the approach, and implemented it in five test States in 1981, calling it UI Random Audit. When initial results of RA confirmed what the six-city study had found, other States were added until RA operated on a voluntary basis in 46 States in 1984. As a result of concerns about trust fund solvency, DOL established an interagency committee to review the findings from RA and other systems used for measuring and correcting payment errors in the UI system. This committee recommended that the Department establish a quality control program throughout the national UI system.

Thus, an extensive effort was undertaken in 1984 to ensure that all the needed data was collected during the investigative process so that it would be available for use in analyzing the origin of inaccurate payments as well as other critical aspects about the individual claim. States were invited to identify all types of data which could be used for analysis of overpayments and underpayments based on a sample of paid claims. The data

2

collection instrument which resulted from this effort provided State and Federal staff with a means to calculate error rates and to assign the cause and responsibility for those errors. In addition, since an extensive amount of data was collected on all aspects of the individual cases sampled, the data collection instrument provided potential sign posts to point to areas requiring further analysis or study. More than 100 data elements were designated for collection on each sampled case. Since the system was based on a probability sample, projections could be made to the entire intrastate UI population, the universe.

Before proceeding with implementation, then Secretary of Labor William Brock ordered a comprehensive review of the proposed QC design. This review, completed in the fall of 1985, resulted in a consensus agreement on a series of design changes and operating principles to govern program implementation. After adjustments to incorporate these new principles, QC was implemented on a voluntary basis in April 1986. Regulations to govern the program were first published for public comment in July 1986. The final rule was promulgated on September 3, 1987. Benefits QC has been operational in all States and jurisdictions except the Virgin Islands since its "official" inception in October 1987.

During the first two years of the core Benefits Quality Control program, pilot tests were conducted in a number of States on two related UI program areas. One set of tests took place in five States on claimants who were initially denied UI benefits. The other pilot, conducted in Idaho, was designed to gauge the validity of BQC case investigative information gathered through telephone interviews.

Revenue QC (RQC) is currently under development, with eight States conducting pilots to test the ability of the RQC design to evaluate the strength and weaknesses of UI tax operations. These tests will run from July 1991 through July 1992.

3. **SESA Special Studies**. During the period of voluntary implementation of Quality Control (April '86 - September '87), participating SESAs had an opportunity to run the system, identify any operational problems, and correct them before they could adversely affect the program. Realizing the potential importance of such experience, the National Office encouraged SESAs to conduct "special studies" of particular problems or aspects of proposed QC methodology and structure, which was then under review by and pending approval of the Office of Management and Budget. This proposition, announced in UIPL 29-86, anticipated that the results of such special studies could lead to useful modifications of QC data collection or other planned methodology before their incorporation into the Federal regulation.

UIPL 29-86 established ETA policy of authorizing SESAs to deploy QC staff resources into the development and implementation of special studies, so long as each SESA investigated annually a minimum sample of 400 UI weeks compensated. Guidelines for developing these studies are extensive and methodology is rigorous. Options for such studies include modification of sampling methodology (e.g., cluster or stratified sampling), augmented sampling (e.g., increased sample size for claims processed in specific problem offices), and data collection methodology (e.g., alternative data collection via telephone interviews of claimants or employers).

As conceived, special studies conducted by SESAs were potentially important to the National Office in either confirming basic QC methodology as planned, or in providing information upon which selected refinements could be made. Response to the invitation was light; three special study proposals were approved during the first two or so years of BQC. Nevertheless, UIPL 29-86 set an important precedent in National Office policy of allowing SESAs to make use of QC staff resources in studying a wide range of UI program issues and problems. The next step was to encourage SESA managers to tap this resource in undertaking short-term research and program studies which may lead to UI program improvements, and ultimately to payment error reduction. This development is described below.

4. __Emergence of QC Program Improvement Studies__. Not long after mandatory benefits QC was implemented, it became apparent that a rich database would be created in each UI jurisdiction. It was apparent, too, that this data could be of significant effect in the improvement of UI payment operations if it were well analyzed and interpreted. DOL provided the SESAs with data analyst and system support positions to assist State UI directors in the utilization of QC and other UI program data. Subsequently, training was conducted for the QC analysts, data user groups were formed, and a national training center was established (by a grant to the Minnesota agency) to customize courses to meet the needs of SESA staff in the analysis and presentation of UI program data to management.

In December 1987 the Department issued UIPL 9-88, a sequel to UIPL 29-86, that confirmed the option of the SESAs to divert QC staff resources from sampling, with National Office approval, to conducting program studies. This action underscored ETA policy of aiding State agencies undertake serious analysis of UI program data as a means of translating general findings into practical program improvements*.

Work in Region X with State UI staff in early 1988 produced a PI study model, called the Quality Improvement Project (QIP). Several SESAs later adopted the QIP model for planning and conducting QC-based studies. Others devised their own study plans and methodologies. To facilitate an exchange among the SESAs of findings and results of such studies, as well as information on other UI program improvement actions, the National Office established the UI/QC Program Improvement Clearinghouse (via UIPL 45-88) in June 1988.

During the first year following issuance of UIPL 9-88, twelve States undertook PI studies subsequent to National Office approval of their study proposals. By mid-1991 half of the States had undertaken one or more approved PI studies, and had generated more than 100 final study reports. These studies have covered a wide range of UI program issues and problems. The scope and quality of the study designs, and the results achieved, vary significantly. A number of these studies reveal the work of staff experienced in research study design. In other cases the PI study proposals and/or final reports suggest that a number of the SESAs involved may not now be able to marshal such expertise, due in some measure to reductions in Research and Statistics UI staff units over the past few years.

---

*Program improvement is used here to cover a wide range of operational, policy, and program changes that may be undertaken by SESA management to ensure proper operation of the UI payments system.

Collectively, this UI study activity among the States indicates commendable effort toward using QC and other UI program data to foster program improvements. At the same time this experience prompted ETA's Office of Quality Control in the Unemployment Insurance Service to launch, in late 1990, a special technical assistance project that would provide technical assistance to SESAs in the design and implementation of UI program studies. This technical assistance manual (TAG) is a major product of that grant project, which was carried out from October 1990 through March 1992 by staff of the Pennsylvania Department of Labor and Industry, Bureau of Unemployment Compensation.

5. **Guidelines for Conducting QC-Supported UI Program Studies**. SESA managers who wish to enlist QC resources (staff) in planning and conducting PI studies must secure prior approval of study proposals from both Regional and National Office UIS staff. Detailed guidance for the preparation and submittal of study proposals is incorporated in a UI program letter (UIPL 37-91) issued to the SESAs in August 1991. This is the third such directive since UIPL 29-86 was published in late 1987. A copy of the current UIPL is included in Appendix A.

Study Proposal Content - Issues and problems proposed for further study must be based at least in part on QC data, and should reflect the concern and expressed needs of SESA decisionmakers. The PI study proposal must address five basic standards spelled out in the guiding UIS directive. These requirements are:

(a) the problem or issue to be studied;
(b) the objectives of the proposed study;
(c) the methodology to be followed;

(d) the results or findings expected and a projection of how to use these findings to achieve UI program refinement;

(e) a workplan of the UI/QC resources to be used in the preparation and conduct of the study and submittal of the final report.

__Timing of Proposal Submissions__ - Inasmuch as approvals of QC staff diversions to PI studies are often related to the Fiscal Year (FY), it is highly desirable that SESAs submit completed study proposals to the Regional Offices at least two months prior to the anticipated start date of a given study. The Regions, in turn, will transmit suitable proposals, along with their respective assessments and recommendations, to the National Office for review and approval.

Approval of program studies is not necessary whenever the State maintains its required QC weekly sampling and funds the project with regular UI resources. Advance approval from the National Office is required if a State plans to fund the study by reducing its required annual QC sample, and deploys QC staff in the undertaking. No SESA may reduce its Quality Control sample below 400 cases per year.

6. __Use of the Technical Assistance Guide (TAG)__. The materials contained in this TAG are intended to supplement and clarify how a State can utilize program improvement studies as a vehicle for continuing improvements of the UI program. It is not possible to create a manual that can be all things to all people. The TAG, as designed, is intended to serve SESA personnel in planning for a study based upon findings from QC--and other UI program data--that identify potential problems. Most of the TAG material is

self-explanatory and is based on the premise that basic knowledge of statistics is available within the agency. In some SESAs it may be necessary to reach outside the agency to other community resources.

In its promotion of UI program improvement studies, the Department is not advocating rigorous controlled experiments for knowledge development that does not have practical applications. Studies designed using this TAG should be action oriented toward fulfilling management's need to resolve operational problems. As you read through the materials, consider what your data tells you and think in terms of what needs to be done to serve the customers of the UI system.

BLANK PAGE

CHAPTER II

PROBLEM IDENTIFICATION

Checklist

1. Have all relevant data sources been analyzed?

    a. ____ Is the problem reflected in QC data/findings?

    b. ____ Can other UI data sources be used to assist in defining the problem?

    c. ____ Have the limitations of the data sources been considered?

    d. ____ Is additional data collection needed?

2. Have all relevant parties been involved in the problem identification process?

    a. ____ Have appropriate UI technical sections been consulted?

    b. ____ Does the UI Director support the problem identification efforts?

    c. ____ Have other affected principals been consulted: Benefits Chief, Field Office Director?

3. Has the problem been sufficiently identified?

    a. ____ Can the problem be attributed to a specific process, procedure, group or cause?

    b. ____ Have the costs to the UI trust fund been estimated?

    c. ____ Have the operational and administrative costs of the problem been estimated?

    d. ____ Given the likely cost to the system, is the problem worth solving?

11

4. Has a report documenting the problem been prepared for UI management?

   a. ____ Does the analysis clearly explain the problem and its associated causes?

   b. ____ Have the data sources and analysis techniques been accurately summarized?

   c. ____ Does the report identify potential solutions?

5. Is the problem solvable?

   a. ____ Can any political/legal barriers be resolved or overcome?

   b. ____ Can any technological barriers be resolved or overcome?

CHAPTER II

PROBLEM IDENTIFICATION

__Introduction__. Problem identification and description is the first step in planning a UI program improvement (PI) study. Problem identification usually involves evaluation of multiple data sources and consultation with several layers of personnel to develop a clear picture of what the "problem" is and to what extent it is affecting UI operations. In some cases the problem identification process may provide sufficient information regarding a given problem and possible solutions to eliminate the need for further study. That is, the problem identification process may result in a decision to directly implement a corrective action or test alternative solutions using resources other than those available for PI studies.

The conclusion of the problem identification process usually involves a preliminary assessment of the need for a program Improvement study. Accordingly, the problem identification process usually involves: (1) clearly defining the area for potential study by utilizing all relevant data sources; (2) including all relevant parties in the process; (3) assessing the cause and impact of the problem on the UI program; (4) reaching a consensus among UI management and appropriate personnel on the problem's solvability; and, (5) preparing a report that summarizes the problem for management. The typical PI study that emerges from this process is likely to focus either on further analysis of existing data, collecting additional data or testing possible solutions to a problem. Each of the above problem identification steps is dealt with in a separate section in this chapter.

Before addressing these essential steps in identifying a suitable problem for PI study, an example of a hypothetical UI study is sketched out briefly. This study example is followed throughout the TAG in presenting and discussing the many issues and decisions necessary for the development of a sound study design and implementation plan.

A. UI Study Example

1. **Background**. Each state's UI program has provisions for claimants to receive UI benefits while employed less than full-time. These earnings must be reported by claimants and, subject to certain minimums, these earnings generally are deductible from a claimant's weekly benefit allowance. (In the BQC system, these earnings are referred to as benefit year earnings, and the reduced weekly benefit payments are referred to as partial payments.) In many States, payment errors associated with partial payments constitute a sizable part of all weeks and dollars with payment errors (including both overpayments and underpayments). Nationwide estimates indicate that benefit year earnings errors account for about 40 percent of all weeks paid in error and about 20 percent of all dollars erroneously paid, even though partial payments represent less than 25 percent of all UI payments.

To reduce benefit year earnings (BYE) errors, some states verify earnings with employers prior to releasing a partial payment. While mailing a verification form to employers is one method of verifying benefit year earnings, a variety of other approaches may be practical. For example, it may be more efficient to require claimants to submit a statement of earnings provided by the employer when signing for benefits.

2. **Study Example**. As a means of relating the TAG to a specific UI area, a hypothetical study designed to evaluate two methods of verifying benefit year earnings will be outlined and used in subsequent chapters. The example study will be designed to compare the effectiveness of a mail verification procedure (the regular procedure) to an alternative method in which claimants are required to obtain earnings statements from employers prior to signing for benefits (the test procedure).

Under the mail verification procedure, the claimant reports partial earnings when the claim is filed, and the agency mails a form to the employer. The employer has five (5) working days to return the form with a breakdown of earnings and missed hours for the claim week. If the earnings report is not returned within the allotted time, the payment is made based upon the claimant's estimate of earnings. Because employers frequently fail to return the form in the required time period, benefits are frequently paid based solely upon the claimant's estimate of earnings. When the earnings reported by the claimant are not correct, a payment error occurs.

Under the alternative test procedure, partial earnings forms will be distributed to some or all covered employers, and claimants will be required to bring a completed earnings verification form, signed by the employer, to the local office when filing a partial claim. The presumption or hypothesis is that the proposed procedure will reduce payment errors because it will eliminate payments based upon claimants' earnings estimates. This conclusion was reached after extensive analysis of various data sources (see sample BYE Proposal, Appendix B). The example study is designed to obtain the results needed to determine whether the test procedure should replace the regular procedure. However, the issues that arise in developing and conducting such a study are relevant for most PI studies. Thus, the general issues involved in the example study are what is important about the example, not the details of the example study itself. Consequently, many details of the example study are not developed in the TAG itself.

B. Utilizing Existing UI Data

1. **Quality Control Data**. As specified in National Office study guidelines, the area of a PI study must be identified as a potential problem by BQC data. Although this places general limitations on acceptable study topics, the requirement is not as restrictive as it may appear. QC data may be used in a variety of ways to determine the need for study in specific areas, and an understanding of the various methods of using QC data may help clarify how study topics can be identified by using BQC data.

Quality Control data collection techniques produce valuable information because the BQC cases are random samples drawn from the UI payment population. Because data from random samples may be generalized to the population with a known degree of precision, BQC error estimates are an appropriate starting point for exploring UI problems. In many instances, however, there may not be sufficient BQC data available to produce reliable estimates for an actual error cause (e.g., misreported earnings). This occurs because the six "error causes" reflected in BQC data are highly aggregated error categories. Taking corrective action usually requires more specific information than that represented in BQC error categories, so comprehensive case reviews or analysis of other data sources is frequently required to determine specific error causes.

16

Quality Control Data Sufficient - When a large number of QC cases are in error due to a specific cause, it may be possible to precisely identify a problem and to generate QC estimates which are reliable indicators of the impact of the problem upon the UI system. In these instances, QC data defines the problem to an extent that makes it unnecessary to review other data sources or collect additional data. In general, additional data collection is unnecessary if the problem can be attributed specifically to a single process, procedure or group (or to a well-defined and small number of processes, procedures or groups). In situations such as these, the SESA may want to move directly to instituting new procedures or testing solutions to a problem.

Quality Control Data Insufficient - Additional data analysis or collection is needed when QC error estimates are imprecise and appropriate corrective actions cannot be determined solely from BQC error rate data. In these situations, it may be necessary to conduct a descriptive study to obtain the information needed to further define the problem. One method of doing this is simply to augment the current BQC sample. By collecting additional data, exploratory data analysis and quality management tools can be applied to identify the sources and causes of process errors.

Other Quality Control Based Indicators - There are other indicators in the QC process that may provide the foundation for a PI study. For example, it may become apparent during QC investigations that a problem exists in the manner in which a particular UI procedure has been designed or implemented. Also, data accumulated in the QC state-optional fields often is useful for problem identification purposes. In any case, some specific ways to analyze BQC data as part of the problem identification process for the example study are explained below.

Partial Payment Study - Quality control data can be used to justify the partial payment study, but it is necessary to consider more than the overall benefit year earnings error rate (cause codes 100-150). Because the study will focus on the reporting of benefit year earnings, an accurate assessment of the problem begins with an analysis of error cause codes 110 and 130. These cause codes relate to over/underpayments caused by the misreporting or misrecording of earnings (110) and vacation pay (130). The remaining cause codes relate to unreported earnings, severance pay, pensions and other causes and are not relevant to the reporting procedure being investigated.

If a State has enhanced its BQC coding scheme to distinguish between errors due to misrecording and misreporting of earnings, then estimates of the earnings reporting problem may be possible. This coding enhancement may be accomplished by assigning numeric codes utilizing the third digit State option field provided in the QC coding structure. If a State has not enhanced the QC coding scheme, however, the error cases with cause codes 110 or 130 will be reflective of a variety of different types of error, and case-by-case review will be necessary to determine the number of cases in error due to misreported earnings and vacation pay.

As another means of using QC data to estimate the magnitude of the earnings reporting problem, it may be helpful to identify all cases with earnings reported during the key week, and to determine the percentage of these earnings reports found to be incorrect during the QC investigation. Included in this analysis should be cases in which the amount was reported incorrectly, but the amount earned was too small to result in a mispayment. These cases are included because, even though they didn't result in a mispayment, they are indicative of a breakdown in the reporting process.

2. **Secondary Data Sources**. There are a considerable number of secondary data sources within each SESA. Some stem from federal reporting requirements, while others are developed within SESAs to satisfy specific informational needs. Reviewing these multiple sources of information is critical when assessing the need for a program improvement study. Although a review of all available data sources may be tedious, it is important because these data sources frequently provide considerable insight into the causes and impact of problems within a SESA's UI operation.

When planning a review of UI data, an initial decision must be made regarding secondary data sources which will provide relevant information. It is important to consider all possible secondary data sources at this point, and failure to do so may result in an unclear description of the problem, incomplete study objectives, or an inappropriate study design.

Major sources of other UI data include, but are not limited to, the following:

UI Research - This area includes all information collected and reported to the USDOL. In addition, states gather information for UI system management purposes which may be available.

Benefit Payment Control - If states gather information to manage/control a specific BPC system (i.e., model crossmatch/recovery system), this information can be a rich source of diverse UI data.

UI Technical Staff - All UI staff from the local office level to the state technical staff may have insight into a suspected problem through system, employer and claimant feedback.

Field Review - This includes information gathered by audit groups and state procedural review teams.

Automated UI Systems - Automated UI systems collect huge amounts of data regarding the filing and payment process. Much of that data is maintained in the system but not reported unless requested by UI staff.

Quality Appraisal - Quality appraisal data is generally available in states and is a good source of data inasmuch as it is gathered as part of a quality evaluation.

3. **Data Source Matrix**. Several secondary data sources provide information related to potential PI studies. In many cases, the available secondary data varies in terms of both reliability and the insight it provides regarding a problem. Thus, it is helpful to construct a data source matrix to guide the review of various data sources. The data sources that might be used to evaluate the example BYE study illustrate the general idea. Table II-1, page II-21, provides a data matrix that identifies data sources that may be used to define/identify problem areas in the UI partial payment process. An "X" indicates that the data source is relevant to the benefit year earnings problem.

The BQC database may be useful for identifying specific causes of BYE errors. In addition to the standard QC error reports, intensive case-by-case review of BYE cases may produce additional insight into the cause of the problem.

Benefit Payment Control (BPC) overpayment reports provide valuable information regarding mispayment detection points and the scope of potential problems. Many SESAs have refined the recording of BPC overpayment data in a manner that makes it

20

possible to select and analyze cases by specific cause. This is helpful when attempting to establish which portion of a current procedure is causing errors. In the example study, BPC data can be useful if the BPC coding scheme has been modified to allow selection of cases overpaid because of misreported earnings.

The UI Research unit also has data that relates to BYE payment errors. The research database has information regarding the accuracy and rate of occurrence of countable earnings determinations as well as the accompanying overpayment information. These data can provide insight into how well local offices handle earnings cases and the rate of reversal upon appeal.

Local office reviews also gather data that may shed light on aspects of all the processes involved in BYE payment errors. Thus, these reviews are used in the example study to help determine the scope and nature of the problem.

The UI Technical Staff unit has persons with expertise on most PI study issues. In this example, these persons can usually be especially helpful in evaluating nonmonetary issues and determination accuracy as they relate to partial earnings problems.

The above discussion illustrates how a data source matrix can contribute to the problem identification process. Because it would be impractical to create a "master matrix" which details all data sources and the information they contain, it is necessary to prepare such charts on an ad hoc basis. In using various data sources in the problem identification process, however, it is important to recognize some general differences between BQC data and most other UI data sources. These differences are discussed below.

4. **Limitations of Nonrandom Samples**. There are many situations in which BQC data does not provide reliable estimates of specific error causes. In most States, the error rates for payments to school employees and students are examples of error causes that are not reliably estimated with BQC data because too few students and school employees are included in BQC data. In such instances, a study may be useful for collecting additional data to more precisely identify/define the problem. Before deciding whether such a study is needed, however, it is necessary to examine the other UI data sources specified in one's data source matrix (Table II-1). Although review of other UI data may not provide sufficient information to preclude the need for a data gathering PI study, checking such data sources will help determine the appropriate study type.

Secondary Data Limitations - When using alternative data sources to clarify a problem indicated by BQC data, it is important to recognize that there is a distinct difference between BQC data and most other UI data. BQC data comes from a systematic sample that represents the population of all UI payments, whereas most other sources of relevant data (e.g., BPC) apply only to specific groups within the UI system. Using data from such sources can be an important part of problem identification, but findings for specific groups cannot be generalized to the UI population. Once a specific group has been established (e.g., claimants with earnings during 10 weeks and UI benefits during 6 weeks of a quarter), direct inferences can be made only about that group, and those inferences usually cannot be extended to other groups of UI claimants. In fact, mistakenly extrapolating results for special groups (e.g., model crossmatch findings and associated overpayment dollar amounts) to the entire UI population can be very misleading. This does not mean that information from sources other than BQC are flawed or inaccurate. But such data must be carefully interpreted and generally cannot be used to directly construct UI population payment error estimates.

C. Building Organizational Support

1. **Importance of Broad UI Support**. Studies have shown that people in organizations tend to resist changes to their work and work environment.[1] PI studies are by nature a vehicle of change since their outcomes may result in UI process change(s). In order to lessen those effects and the resistance to a study, it is important to coordinate the problem identification effort with all potentially affected agency personnel. In fact, lack of sufficient coordination within the SESA can be one of the most serious problems in successfully completing a PI study and having action taken on study findings. The relevant personnel differ at each stage of the PI project. There are few firm rules to follow, except to involve everyone who has a stake in the study development process and outcome. Failure to include key staff in the formative stages of a PI study often makes it difficult to obtain their support for any recommendations that result from the study. Program improvement studies conducted in a vacuum will usually generate little support for program improvement recommendations. Because the ultimate goal of all PI efforts is to improve operations, it is essential to make every effort to involve all relevant UI staff. Obvious staff to involve in most problem identification processes include the UI Director, Benefits Chief and responsible program area managers. These key staff should be involved in the problem identification process and should provide input while the impact of a given problem on the UI system is being determined.

Generally, the following persons should be considered when formalizing a strategy for developing and maintaining organizational support: (a) SESA Administrator(s), (b) program managers, (c) local office staff and study workers, (d) program improvement Study staff, and (e) outside interested parties

Organizational Coordination - To assist in determining which personnel to involve at various points during the study process, a matrix similar to the one found in Table II-2, page II-22, should be prepared at the outset of the problem identification process. This table should be updated as required through all stages of the PI study. It contains the names or titles of interested staff, a brief description of their relation to the study area, statements regarding their personal concerns, stages of the study during which close coordination is necessary, and a section regarding how and when these individuals will be provided with updates regarding study progress.

Although the key staff mentioned above likely will be involved in most PI studies, it is more difficult to determine when to involve personnel from areas that have less obvious interest in the PI study. This is difficult because of the changing nature of the study environment. Early stages of problem identification may require input, direction and coordination with relatively few key staff. Selecting the study approach and determining data collection techniques, however, require coordination with additional staff. Implementing study findings, may involve yet a different group of staff. Thus, starting with the problem identification process, it is important to ensure that potentially interested parties are included.

Flowchart Approach - After the support of key administrative staff has been secured, a detailed flow chart of the function or process to be studied is useful for understanding both the study issue and who should be considered a primary or secondary party at the various stages of the PI project. Flow charting the study function/process has many additional benefits, such as helping to identify critical input, processing and output points that significantly impact the problem.

Primary vs. Secondary Parties - The fact that there are primary and secondary parties involved, and that these individuals can change during the study merits additional comment. Primary parties may be defined as those individuals: (a) whose input is required to develop study plans (e.g., personnel from QC, BPC, Field Audits, Tax Audit Data); (b) who will be required to participate by changing their work habits (e.g., local office claims interviewers) and; (c) who likely would be required to execute new methods of doing business as a result of the study (e.g., central office operational staff, ADP staff).

Secondary parties may be described as those individuals or groups that will not be directly affected by the study process or its potential outcome, but may be inconvenienced by the study or have an interest in its outcome. An example may be a study of base period wage reporting errors in which questionnaires, field audits or other contacts with the employer community are required. It may not be necessary to closely coordinate these activities with the Director of Tax Operations and his/her field auditors, but informing them about the study is important, and it prepares them for any inquiries regarding study activities made by the employer community. These secondary individuals or groups are difficult to identify at times, but careful preparation of process flow charts with potentially affected or interested parties specified at each step helps prevent major coordination problems.

Forward Looking Approach - The process of obtaining/maintaining organizational support must be forward looking because it begins when a PI study is planned and continues through the completion of the study. This means that, during the problem identification stage, it is useful to consider organizational support issues that will be involved throughout the planning period and the implementation of the study, and even during the development of the final report. Some useful tips for doing this are summarized in Table II-3, pages II-23 and II-24.

D. Problem Cause and Impacts

1. **Estimating Costs**. After examining essential data sources and involving appropriate staff, it should be clear whether the purpose of the PI study is to evaluate more data in order to isolate the cause of the problem or evaluate whether a change in a specific process or procedure will help solve the problem. Regardless of the study purpose, rough estimates of the dollar impact of the problem are required. When estimating costs, it is important to consider: (a) costs to the UI trust fund; (b) operational and administrative costs; and (c) costs to claimants and employers.

Trust Fund Costs - Although QC data may provide an indication of trust fund costs, when estimating these costs it is important to note that corrective actions, regardless of form, that result in the elimination of overpayments may not amount to any UI trust fund savings. At first glance, it may seem inconsistent that eliminating overpayments does not result in a savings. However, corrective actions often are intended to result in higher claimant compliance. For example, a better BRI may result in more claimants adhering to the active worksearch requirement. Assuming these claimants subsequently draw the same total benefits while in compliance as they did while out of compliance, the draw on the UI trust fund would remain constant. In fact, assuming that some dollars were recovered from claimants overpaid due to worksearch, there actually could be a net loss of trust fund dollars from increased claimant compliance. The point is that if a potential corrective action relies upon increasing claimant compliance or eliminating a particular rule, the effect need not result in readily identifiable UI trust fund savings.

For the partial payments example used in the TAG, costs to the UI trust fund can be estimated with QC data. In assessing trust fund costs for the study example, however, caution is required to ensure that error estimates are calculated only for overpayments caused by misreporting of earnings, missed hours or vacation pay.

2. Operational and Administrative Costs. To assess the magnitude of operational and administrative costs associated with the earnings misreporting problem, it is necessary to determine the resources required to process an over/underpayment caused by misreported earnings. Based upon this estimate, a general estimate of impact can be obtained. Calculation of these costs is an important factor when determining the benefits that can accrue from a PI study. The following is a purely hypothetical illustration of the impact that reduced processing of mispayments may have on a SESA's UI operations.

Assuming that a SESA has calculated the average per unit cost of detection, adjudication, collection and maintenance of benefit year earnings mispayments to be $100.00, and such mispayments are detected at a rate of 40,000 per year, a potential administrative savings of one million dollars would exist with a 25 percent reduction of such mispayments. Such a reduction could amount to

27

40 positions per year when calculated at a per position funding rate of $25,000.

3. **Claimant and Employer Costs**. Determining costs to claimants and employers may be more difficult, but some consideration should be given to these costs. In this example, claimant costs will be related to actual dollars lost and to the costs imposed on them by the test procedure. These include underpayments or reduced benefits as overpayments are recouped from subsequent checks. For employers, the costs are likely limited to personnel costs related to time spent completing earnings verification reports under either approach.

E. Assessing Solvability

Potential Barriers. Various barriers may adversely affect attempts to resolve any problem addressed in a PI study. Although problem identification is not the study stage where a final assessment of solvability occurs, some assessment is needed to avoid a study that pursues a change that cannot be implemented. In this light, it is useful to consider the following aspects of solvability prior to moving beyond the problem identification stage of a potential PI study:

- -- **organizational barriers** - does responsibility for action to solve the problem fall outside the realm of SESA Administration?

- -- **legal barriers** - will the possible solutions require "impossible" changes in SESA UI law?

28

-- **technological barriers** - can the likely solution be supported by current resources, including EDP resources, or will a large expenditure on new resources be necessary?

-- **administrative barriers** - are potential solutions likely to be administratively feasible?

-- **expertise barriers** - does the SESA have sufficient expertise available to pursue the likely course of the project?

-- **atmospheric barriers** - does a positive attitude toward change exist among the potentially affected staff?

Making an initial evaluation of the potential effect of the above barriers is useful in deciding on the feasibility of a PI project. This evaluation also helps avoid possible pitfalls that could occur as the PI process moves forward.

F. Problem Identification Report

Study Justification. Organizing material into a report that documents findings and conclusions is the last step of the problem identification process. Documenting the problem in a report for UI management should not be overly complicated. Knowing the target audience is helpful in determining the best form and style of the report. Interesting reports that are brief and to the point usually work best. Tables containing raw data should be tucked away as attachments, available to those who are interested in such items.

On many occasions a problem identification report will be scanned only by senior administrative staff before being forwarded to technical staff with specific program area responsibility. This makes the executive summary an extremely important part of the report. The executive summary should not be more than a few pages long and should, if possible, contain graphic illustrations of the major findings. If protocol permits, the executive summary can be reduced to a fact sheet and provided to all interested parties as a method of initial contact.

Several hints may be helpful for preparing the problem identification report:

- Include an executive summary.
- Make the report clear and concise.
- Make the report readable - avoid bureaucratese.
- Ensure graphs are clear and well labeled.
- Produce an honest report that shows the possible costs and pitfalls, as well as the potential benefits.

Keep in mind that one outcome of the problem identification process may be a decision that a formal PI study is not required. The intensive data review and analysis prior to proposing a PI project is intended to clarify: (1) whether a project is required; and (2) what direction the study will take, given what is known about the problem from the available UI information sources.

TABLE II-1

PROBLEM IDENTIFICATION DATA SOURCE MATRIX

UI PAYMENT ACTIVITY	QC	BPC	UI RESEARCH DATA	L.O. REVIEW	UI TECH. STAFF
Non-Mon Accuracy	X		X	X	X
Non-Mon Issues	X		X	X	X
Overpayment Data	X	X	X	X	
Earnings Status	X	X		X	

TABLE II-2*

ORGANIZATION COORDINATION MATRIX

Name of Individual or Group	Relationship to Area of Study	Primary Concerns	Point(s) of Study Coordination	Reporting Schedule
UI Director	Project Manager	All data sources utilized. Proper methodology. Efficient data collection. Reliable study findings.	All phases of study.	Verbal reports weekly. Written reports monthly.
UI Claims Manager	Program Manager study area	Accurate assessment of current procedure. Coordination with field staff. Smooth transition to test procedures.	Problem identification. Selecting methodology. Operation of test procedure.	Verbal reports weekly. Written reports as requested.
BPC Chief	Program Area Manager	Proper use of BPC data to identify problem.	Problem identification. Assessment of test procedure. Impact on mispayments.	Verbal reports as requested.

*Table II-2 adapted from an instrument developed by Phyllis Jacobsen found in "How to Communicate Evaluation Findings" by Morris, Fitz-Gibbon Freeman. Sage, 1987. page 18.

TABLE II-3

ORGANIZATIONAL SUPPORT CONSIDERATIONS

Problem Identification and Study Objectives

- Be open and specific regarding data assembly during the problem identification phase.
- Openly communicate the reasons for analyzing the problem area.
- Involve program area staff in analysis and problem identification.
- Make key staff aware of problem identification findings.
- Solicit input regarding the development of feasible study objectives.

Study Approach and Operational Procedures

- Solicit input from technical and operational staff on the study approach.
- Provide key staff with an honest evaluation of study duration and expected level of staff involvement.
- Incorporate worthwhile input into study design.
- Solicit input and technical assistance from agency, Regional and National office experts.
- Be realistic concerning the possible negative outcomes of the study.
- Provide sufficient training to the operational staff before study begins.

Study Operational Period

- Provide frequent progress updates to key staff.
- Institute suggested changes that arise during the study, if they do not affect study reliability or credibility.
- Make on-site visits to evaluate the success of the project.
- Provide adequate notice of study procedural changes and explain why they are needed.
- Create an environment of discovery for everyone involved in the study.

Data Analyses, Conclusions and Recommendations Phase

- Openly discuss the indications suggested by the data.
- Prepare a comprehensive data analysis plan.
- Solicit UI program staff input in interpreting the data.
- Explain the data analysis techniques to be used.
- Involve appropriate staff in preparing benefit/cost estimates.
- Seek input from key staff regarding appropriate conclusions and the feasibility of recommended changes.

Final Report

- Preview a preliminary version of the report and obtain suggested changes from key staff.
- Involve key staff in developing final report, especially the recommendations and conclusions.

FOOTNOTES

[1] Stan Kossen, The Human Side of Organizations, 5th Edition. (New York: Harper Collins, 1991), p. 282.

CHAPTER III

STUDY GOALS AND OBJECTIVES

Checklist

1. Has a thorough understanding of the study area been developed?

    a. ____ Did all relevant staff have the opportunity to provide input in developing study objectives?

    b. ____ Have the major variables and relationships to the study been diagramed?

    c. ____ Have the effects of secondary or environmental variables on the outcome of the study been addressed?

2. What will be learned from the study?

    a. ____ Is the list of research questions manageable?

    b. ____ Is the information needed to answer these questions available?

    c. ____ Is the information sufficient to provide necessary insight to solve the problem?

3. Are specific study objectives established?

    a. ____ Does each objective state only one purpose or aim?

    b. ____ Do the objectives utilize verbs that describe a measurable result expected from the study?

    c. ____ Will answering the research questions achieve these objectives?

4. Do the primary decision makers support the study objectives?

    a. ____ The primary UI manager(s)?

    b. ____ Are other key decisionmakers willing to commit necessary resources?
    c. ____ Has a preliminary listing of tasks and estimated resource requirements been developed?

BLANK PAGE

CHAPTER III

STUDY GOALS AND OBJECTIVES

In the problem identification process covered in Chapter II (related to BYE methods), important issues were explored that frame the scope of a typical PI study. At this point, it is necessary to identify the purpose of the study, relevant study variables, and understand how these variables relate to the purpose of the study. This area is often referred to as establishing a study's system boundaries or conceptual framework.

1. **Study Purpose**. As an initial step in focusing the study, decisions must be made regarding the overall purpose of the study. These decisions are appropriately made and approved by all relevant UI personnel after reviewing the findings of problem identification effort. This will ensure that the study is designed to provide the information that managers need for subsequent decisions. For the BYE example, it was determined that a study to test a new procedure for reporting partial earnings will be conducted. Any cause and effect relationships between the reporting procedures and the incidence of error will be investigated as part of the study. An example of a statement of purpose/goals associated with this study is:

> To determine the effectiveness of a modified partial earnings reporting method and to obtain information to help determine whether a procedural change is warranted.

This statement is sufficient to establish the general purpose of the study, but it is too vague to serve as the study's objectives. Study goals generally are broad statements which reflect overall purpose. Objectives are more specific statements which concretely define study activities and specify the desired end results of the study.

It is easy to see that there are some major implications for the study approach inherent in the above goal statement, however. Specifically, "determine the effectiveness" immediately points to an attempt to establish causation. This requires evaluating the degree to which a procedure "causes" proper or improper reporting of earnings. This effort to measure cause-effect points to a study approach that utilizes either "experimental" or "quasi-experimental" designs to compare the effectiveness of different procedures. (These designs are discussed in Chapter IV).

2. System Boundaries. Before proceeding to the point of specifying objectives, it is important to take a step back and consider the overall system within which the study will be conducted. If not already done, this will require the study planner to meet with relevant UI personnel to discuss the general purpose of the study and obtain their views on the appropriate objectives for the study.

To assist in the development of achievable objectives, it is recommended that the study planners focus their discussions utilizing a two step process: (a) the development of a research diagram which identifies relevant influencing variables, and (b) development of research questions which establish the relationship between these variables. After this is done specific study objectives for the project can be developed.

Study Variables - Regardless of whether additional data collection or the effect of a policy or procedure change is the study focus, it is important to consider the variables and assumed relationships involved in the study. The underlying assumption, or hypothesis, for the example used is that the test procedure will result in more accurate reporting of partial earnings than the existing procedure (for reasoning, see BYE proposal, Appendix B). This immediately identifies the key variables and relationships for the example PI study. But it also is important to consider other influences and variables that may affect study outcomes. It is necessary to identify these other variables/influences in order to adequately control for their effects in designing the study.

To this end, development of a diagram/chart which identifies these variables is an essential step in framing the boundaries of the study. Chart III-1, page III-8, is an example of how such a diagram might look for this example. The chart, developed in conjunction with interested agency personnel, identifies both the variables to be collected and establishes their relationship to the study's primary cause and outcome variables.

Research Questions - The primary purpose of any PI study is to obtain information to help UI managers evaluate existing or proposed policies, processes or procedures. The PI study objectives will frame what information will be used in the evaluation. The second step in establishing clear study boundaries is development of research questions based upon the assumed relationships demonstrated in the research diagram.

Many questions may be identified for a partial earnings PI study. The general question used in the example is: how effective will a new partial earnings reporting procedure be? This question may be broken down into the following list of associated questions that would make it possible both to compare the effectiveness of the two reporting procedures and also to make an informed decision regarding the implications for the UI population.

a. What proportion of sampled cases using the current reporting method have correct partial earnings? How do these sample estimates project to the population?

b. What proportion of sampled cases using the test reporting procedure have correct partial earnings? How do these sample estimates project to the population?

c. Are the earnings of the test procedure sample reported correctly at a higher rate than those of the current procedure sample?

d. What is the estimated impact of implementing the test procedure statewide?

e. How are significantly associated variables affecting the incorrect reporting of earnings for the two verification procedures?

3. Study Objectives. The next essential task in establishing system boundaries is to convert each of the above research questions into clear, concise objectives. Objectives should be viewed as specific statements that define study goals as precisely as possible. Several techniques helpful for writing useful objectives are:

a. State only one purpose or aim per objective. Most PI studies will have multiple objectives, of course, but separate objectives should be stated for each purpose.

b. Specify a single end purpose or expected result per objective.

c. Use verbs that describe an observable or quantifiable result.

The key study outcome variable for the BYE study is the relative effectiveness of the two reporting procedures in reducing partial earnings payment errors. As the study design is refined, it will be necessary to develop an objective, quantifiable measure for assessing the key study outcome variable as well as related variables identified in the research diagram. For instance, the proportion of audited claim weeks in which the LO received the correct information from the claimant vs. the employer might be a quantifiable indicator of the relative effectiveness of the two reporting procedures.

By using the above process, specific objectives for the example BYE study were written. The four objectives projected for the study are:

a. To assess the accuracy of the current benefit year earnings verification procedure by selecting a systematic random sample of partial payments.

b. To assess the accuracy of a modified benefit year earnings verification procedure by selecting a systematic random sample of partial payments.

c. To compare relative accuracies of the current and modified benefit year earnings verification procedures.

d. To implement the test procedure if feasible.

4. **Study Feasibility Overview**. Once the primary study objectives have been developed it is necessary to make a preliminary assessment of required and available resources in the planning phase of a PI study. Before the entire study is planned, the estimate of required resources necessarily is somewhat sketchy, but more detailed estimates are possible as the study methodology is refined. The fact that initial estimates are sketchy does not minimize the importance of evaluating resources early in the study planning stage. In fact, it is beneficial to construct a preliminary resource estimate during the planning phase for at least two reasons. First, a preliminary resource estimate helps avoid a study that cannot achieve its objectives, given likely resources. Second, it saves time and effort in preparing the final workplan for the PI proposal. The following are several basic steps that may be used in estimating required resources:

a. Prepare a sequential listing of evaluation tasks to be performed and products to be produced.

b. Allocate a specific time allotment for each task.

c. Estimate the staff and other resources needed for each task.

d. Estimate the quantity or amount of total resources required[1].

U.S. Department of Labor
Employment and Training Administration
Washington, D.C. 20210

CLASSIFICATION
UI

CORRESPONDENCE SYMBOL
TEUQI

DATE
June 15, 1992

DIRECTIVE : UNEMPLOYMENT INSURANCE PROGRAM LETTER NO. 31-92

TO : ALL STATE EMPLOYMENT SECURITY AGENCIES

FROM : DONALD J. KULICK DJ Kulick
Administrator
for Regional Management

SUBJECT : Transmittal of A Guide for Planning UI Program Improvement Studies

1. Purpose. To provide State employment security agencies (SESAs) with copies of a newly published guide for planning Unemployment Insurance (UI) program improvement studies.

2. References. UIPL 37-91 (August 13, 1991); and UIPL 7-91 (November 21, 1990).

3. Background. In September 1990, the Department of Labor awarded a grant to the Pennsylvania Department of Labor and Industry to conduct a technical assistance (TA) project to aid SESAs in planning and implementing UI program improvement studies. This project was carried out over an 18-month period by staff of the Pennsylvania Bureau of Unemployment Compensation. During this period, project staff surveyed SESA principals regarding their needs for and interest in UI program research technical assistance, planned and conducted three research design training workshops primarily for staff of the State agencies, and developed a Technical Assistance Guide (TAG) for use by SESAs in designing and conducting UI program improvement studies. Copies (four) of this document, entitled A Guide for Planning UI Program Improvement Studies, accompany this program letter. Additional copies may be obtained upon request.

4. Action Requested. All SESA Administrators are asked to provide a copy of this UIPL and the Guide to the UI Director, the Chief of Benefits, and the person who has primary responsibility for the UI Quality Control program.

RESCISSIONS

EXPIRATION DATE
June 30, 1993

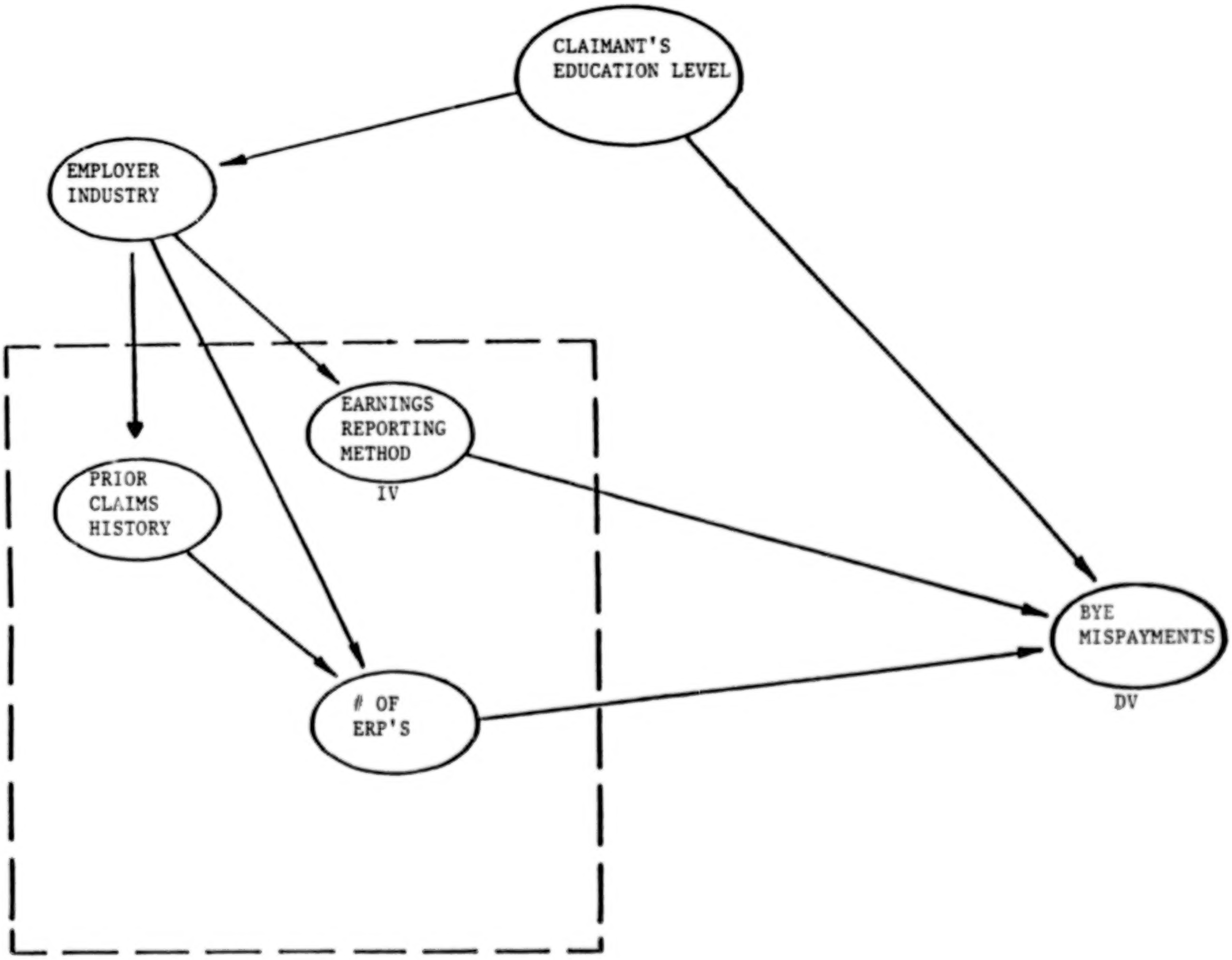
DISTRIBUTION

VL

5. Inquiries. Direct inquiries to the appropriate Regional Office.

6. Attachment. A Guide for Planning UI Program Improvement Studies (four copies).

CHART III-1
BYE
STUDY VARIABLE
DIAGRAM



*adapted from Poister, T.H., Public Program Analysis: Applied Research Methods, page 43.

FOOTNOTES

[1] A.B. Blalock, Evaluating Social Programs at the State and Local Level. Kalamazoo, Michigan: Upjohn Institute, 1990, p. 308.

CHAPTER IV

STUDY APPROACH

Checklist

1. Is the type of study needed to answer the relevant policy questions or deal with the identified problem carefully specified?

   a. ____ Are study purposes clearly defined?

   b. ____ Have nonexperimental (descriptive, exploratory or case study) approaches been considered?

   c. ____ Have more "rigorous" experimental or quasi-experimental studies been considered?

   d. ____ What justifies the study approach selected?

   e. ____ Are study procedures clearly defined so the study could be replicated?

   f. ____ What kind of evidence/results is the manager likely to require before he will take action?

2. Is the study approach reasonable?

   a. ____ Can the study approach answer the relevant research questions?

   b. ____ Were the tradeoffs among realism, randomization, and representativeness considered?

   c. ____ Can the study be completed within the desired time frame?

   d. ____ Are the costs of the study less than its likely benefits?

3. Is the study structured to correctly identify specific study effects or might "external" factors make it difficult to reach valid conclusions? This means that any study must guard against results that are not due to the study itself but instead are due to external factors such as:

a. ____ Uncontrolled external events during the study period (e.g., a large increase in the claims load).

b. ____ Changes that occurred during the study, but not because of the study (e.g., a change in the management structure).

c. ____ Changes in measurements during the study period (e.g., a change in appeals or overpayment counting procedures).

d. ____ Choosing a study group based solely on extreme values (e.g., including only exhaustees).

e. ____ Biases due to incorrect selection of persons for comparison groups (e.g. are study participants volunteers).

f. ____ Differential response or drop-out rates for comparison groups.

g. ____ Artificial study effects not likely to be repeated in an operational setting (e.g., study group behavior changes resulting from the study itself).

h. ____ Changes in respondents that occurred during the study, but not in response to changes in study variables (e.g., a change in claimant composition during the study period).

i. ____ The effects of testing, where the respondents either learned from taking a test or performed better simply because of the test.

j. ____ Changes in respondents that occurred due to "regression toward the mean" (this is only relevant when respondents are selected on the basis of extreme values, such as having exhausted their benefits).

4. Do the study results apply only to the study group or can they be generalized to a larger group for policy purposes?

   a. ____ Is the larger group (population of interest) clearly defined?

   b. ____ Does the study group (sample) accurately represent the larger group (population)?

48

c. ____ Can the treatment be applied to the larger group in exactly the same manner that it was applied to the study group?

d. ____ Have technical considerations involved in "scientifically correct" studies been weighed against the reality of available resources?

e. ____ Could the study be replicated and would the results likely be the same?

BLANK PAGE

CHAPTER IV

STUDY APPROACH

Several decisions are required to identify an appropriate PI study approach. The relevant policy questions must be identified to determine how the proposed study relates to those questions and whether the study can be completed within the required time frame. It also is important to make a preliminary assessment of the likely benefits and costs of the study to be sure the proposed study approach is appropriate. All studies are affected to some degree by external factors that must be controlled or accounted for in some way in order to obtain valid results. Once study results are available, it is important to specify the conditions under which those results may be generalized beyond the study itself for policy purposes. These issues are identified in this chapter, and illustrated below.

1. Validity Threats. Because it cannot be concluded that a change in one variable is responsible for changes in another unless one has ruled out the possibility that the observed change was caused by some factor other than the treatment, it is necessary to consider the following threats to internal validity. They are generally categorized and defined as follows:

a. History - events other than the independent variable that could have affected the dependent variable.

b. Maturation - natural changes taking place in the process being studied.

c. Statistical regression - if samples for study or action are picked because of an extreme position, such as maximum

WBA claimants, the instability of extreme scores may cause the observed change.

d. Selection - the way that cases are selected for a study or a program could affect the way they react to the program. Volunteers, for example, are usually different from those not volunteering.

e. Experimental mortality - cases, particularly people, who drop out of a study before it is completed. If they are systematically different from those who remain, the results will not be valid.

f. Testing - a situation in which the initial measure or test influences the subjects' behavior and this affects the outcome of the posttest. Claimants who score low on a pretest may work very hard because of this and not because of the treatment.

g. Instrumentation - if the measuring instrument (DCI) used to collect the data changes between the beginning of a study and its conclusion, the results may not be valid.

h. Design contamination - a condition occurring when participants know that they are in a study and act differently because of it. If study subjects have an incentive to behave in such a way as to make the program succeed or fail, the design has been contaminated (Hawthorne effect).[1]

Each of these threats to study validity should be given careful consideration when selecting a study design, during the operation and data analysis phases of the study. The effect of these threats can result in findings that are very misleading. Obviously, predication of a program change based upon flawed data could have serious quality implications.

Validity Questions - The following is a list of questions that can be asked by the study planner or administrator when evaluating the possible impact of validity threats on a study that attempts to establish a cause and effect relationship between two variables.

a. Based on history: Could an event other than the independent variable have caused this relationship?

b. Based on maturation: Would a similar change in the dependent (outcome) variable have taken place with the same passage of time without the occurrence of the independent (cause) variable? Are the subjects of study undergoing natural changes?

c. Based on statistical regression: Did subjects exposed to treatment represent extremely high or low categories of the dependent (outome) variable? Were subjects chosen for the study because of their high or low category of the dependent (outcome) variable?

d. Based on selection: Were the people exposed to the independent (cause) variable systematically different from those not exposed? If subjects were volunteers, could this have affected the findings?

e. Based on experimental mortality: Were some people initially exposed to the treatment later assigned to the nontreatment group? Did some people who started the study drop out before finishing?

f. Based on testing: Was a pretest administered that could have affected responses to the posttest?

g. Based on instrumentation: Did the instrument to collect the data (DCI) change between the pretest and the posttest?

h. Based on design contamination: Did the program participants mingle with nonparticipants? Would program participants, nonparticipants or both have an incentive to change their behavior and make the program succeed or fail?[2]

2. **Types of Studies**. Social science research, including UI research, can be divided into four major categories: (a) laboratory experiments; (b) field experiments; (c) field studies; and (d) survey research. This breakdown stems from two sources, the distinction between experimental and nonexperimental research and the distinction between laboratory and field research.

Experimental vs. Nonexperimental Research - Experimental research may involve both experimental control of key variables and random assignment of subjects to treatment and control groups. In contrast, nonexperimental research involves analysis when it is not possible to control the variables of interest, either because the results analyzed already have occurred or because it is not possible or desirable to experimentally control the variables of interest. In nonexperimental research, relationships among variables of interest can be determined only through statistical analysis because the researcher has no direct control in structuring those relationships. Nonexperimental research generally occurs when cost, time or other factors preclude experimental control or random assignment.

Quasi-experimental research involves the introduction of something akin to pure experimental design, either in terms of how data are gathered or in terms of selecting the particular cases for which data are collected. The idea is to gain some of the advantages of an experimental approach in a situation where the researcher cannot fully control the events or process under study. As would be expected, experimental data usually are more difficult to gather than quasi-experimental data, and quasi-experimental data usually are more difficult to gather than nonexperimental data.

Laboratory Research - Laboratory experiments attempt to control variations in all or nearly all of the possible influential variables under investigation. Laboratory experiments also attempt to eliminate the effects of all variables that are not pertinent to the immediate problem under investigation. That is, laboratory experiments are characterized by both experimental isolation and constant conditions. Laboratory experiments obviously are rare in the social sciences and rarer still in the UI system.

Field Research - A field experiment is a research study in a realistic situation in which one or more variables of interest are investigated under as carefully controlled conditions as the study permits. Because all conditions cannot be controlled fully, it is very important to thoroughly describe and illuminate the precise context and conditions involved in field research. Nonetheless, notice that the contrasts between laboratory experiments and field experiments are mostly matters of degree.

Field studies are nonexperimental studies aimed at discovering the relations and interactions among variables of interest. Any studies that systematically pursue relationships among variables, that are nonexperimental, and that are done in real life situations (e.g., schools, factories, business firms or UI local offices) are considered field studies. Case studies are the most common example of field studies because they are relatively easy to conduct. Case studies use multiple sources of evidence to study a relationship of interest, such as whether some UI local offices more effectively carry out certain tasks than others. In the example used in this TAG, a question for a case study might be whether some local offices more effectively or more accurately process partial earnings claims than others.

Field studies, including case studies, can be divided into two broad types, exploratory studies and studies designed to test a particular presumption, hypothesis or theory. Exploratory studies seek to determine whether certain relationships may exist, rather than to test presumptions or hypotheses that are specified in advance. Exploratory studies have three purposes: (1) to discover significant variables in the field situation; (2) to discover relations among variables; and (3) to lay the groundwork for more systematic and rigorous testing of presumptions or hypotheses.[3] An example of an exploratory case study would be an analysis to determine whether processing all partial earnings claims centrally in one office produces fewer errors than having them processed in various local offices.

Survey Research - Survey research is a type of exploratory field study that focuses on the accurate assessment of the characteristics of whole populations of people, usually on the basis of a sample. The goal of survey research is not to measure relations or test hypotheses. There are no "dependent" or "independent" variables and there are no experimental controls in survey research. However, the results of survey research often are analyzed with nonexperimental methods to determine whether some relationships of interest may exist in the data. An example of survey research would be to question a sample of claimants about certain services (e.g., the processing of partial earnings claims) they receive at local offices in order to estimate how these services are viewed by the population of UI claimants.

3. **Selection of Feasible Approaches**. The experimental design that best fits any particular situation involves complex considerations that often exceed the technical expertise available in BQC units. However, some SESAs have persons with strong research skills in QC, LMI or UI research units. Further, the required expertise also is available in universities, other institutions of higher education or private sector firms. In addition, there are numerous college texts that detail experimental design issues (see Appendix A), and these will be useful to those with adequate research background. In any case, these experimental design issues are far beyond the scope of this TAG. An appropriate experimental design is selected and used throughout the remainder of the TAG to illustrate several important points about the partial earnings example. But the various considerations that led to the selection of this design are not elaborated in any detail.

Applying Various Approaches to BYE Example - Several of the study approaches described in the prior section can be ruled out immediately for this problem. For example, a laboratory experiment is not feasible because the new procedure must be tested in a realistic setting. Survey research is not feasible since claimants already are supposed to give their best estimates of partial earnings under the existing procedure. Thus, the obvious approach for evaluating this proposed procedure is to conduct a field test. This illustrates the general principle that a field experiment usually should be considered when a problem is well defined and involves the evaluation of a small number of alternatives. In this instance, the SESA has identified the type of form it wants to use and the procedure it wishes to test. Accordingly, a field experiment that compares payment errors under the current system with payment errors under the proposed system clearly is feasible.

The experimental design appropriate for this test of partial earnings reporting procedures is the non-equivalent control group design. The essence of this proposed design is that it requires both random assignment of persons to both partial earnings reporting procedures but also control over the independent variable (the filing process). Altering the filing process poses no difficulty, but random assignment of claimants to the proposed and existing filing processes is not possible because the partial earnings forms must be distributed to employers, not claimants. Thus the primary sampling units must be firms, and this results in a "cluster sample" of claimants (rather than a simple random sample). Cluster samples can be analyzed, but the analysis is somewhat more complex than analyzing simple random samples. Thus, this example illustrates how to approach issues involved in handling data not generated from a pure random sample. This is useful because pure random samples are extremely easy to analyze, whereas cluster samples are slightly more difficult to analyze but often may be involved in PI studies.

Drawbacks/Considerations - A potential problem with the proposed approach is that employers may object if only some of them are subject to the new procedure. This might be minimized by requiring employers in one geographical area to use the new procedure, and employers in another region to use the existing procedure. However, the procedure selected for this field experiment assumes the following sampling procedure is acceptable in terms of differentially treating claimants and employers for the field experiment. Because using a new employer reporting form involves several costs, including the education of both affected claimants and employers, the experimental procedure is limited to a specific subset of employers (in one region of State).

Requiring the same employers to provide partial earnings information under two different reporting procedures probably would be confusing for them. Thus, the new reporting form will be provided to a sample of all employers in one region of State and specific information on the use of the form will be provided to that set of employers. Then, all claimants filing partial earnings claims that involve any of the employers in that sample will be required to comply with the test reporting procedures during the study period by obtaining the required earnings form from their employers.

Control Group - Although, the test group is made up of all claimants from a "cluster sample" of firms, as discussed above, the control group used for the study can be based on a random sample of the claimants who file partial earnings claims in a separate region of the state (i.e., the control group is just a sample of claimants who use the regular reporting procedure in Region A). However, this means that the control group involves a two stage sampling plan that results from separating employers into two groups -- those that report earnings to the agency under the normal procedure for control group claimants and those that report earnings on the form provided to claimants by their employers under the test procedure -- and then selecting a simple random sample of claimants from those groups of claimants.

__Some Alternative Designs__ - The experimental design sketched in the prior section may be difficult or expensive to implement, so it may be worthwhile to consider a more feasible plan that sometimes is possible with a quasi-experimental approach.

Quasi-experiments are intended to be less expensive or time consuming than true experiments, while sacrificing only some generalizability. A quasi-experimental design that could be used for this study is the time series experiment. Under this approach, the idea is to use the existing procedure for some number of periods, say the first 13 weeks of the study. Then, the new procedure is used for the next and last 13 weeks of the study. In this design, a control group is not used. Instead, measurements over time are tracked to isolate the effects of external time related factors from the effects of the treatment. Then, statistical controls can be used to account for extraneous factors, such as industry mix or claims load changes, during the course of the study (see Chapter VIII). After controlling for these other factors, it then is possible to evaluate whether the new procedure reduced partial payment error rates by a statistically significant (and important) amount. However, this approach offers no advantages over the proposed approach in the prior section, and it has the additional disadvantage of educating (some of) the same employers and claimants to use two different reporting procedures during the study. Thus, this quasi-experimental approach is not used for the TAG example.

__Field Study__ - The effect of the proposed change in the partial earnings filing procedure also could be studied by means of a field study, as opposed to the field experiment explained in the

60

prior section. The potential advantage of a field study, compared to a field experiment, is that the former relies on existing data and this eliminates the need for developing an experiment to generate data. Thus, a field study is an attractive approach to consider. For the TAG example, a consideration would be whether existing BQC data could be analyzed to determine the impact of the new partial earnings reporting procedure. If cases in which the partial earnings reported by claimants and employers differed could be found in past QC cases, an assumption then could be made concerning the extent to which the new procedure would have eliminated the payment errors that resulted.

If the new procedure appears to be worth implementing even under a worst case scenario based on the field study, then an experiment obviously is not necessary. Instead, the decision to implement the new reporting procedure for partial earnings could be made. This approach obviously has strong appeal whenever circumstances justify it. But very strong assumptions normally are required to take action on such limited information as would be available in most States. Thus, even if action is based on such information, it still is important to track results after implementation (probably on a pilot-test basis) to be sure actual results conform adequately with expected results. In any case the field study is not selected for the TAG example because so few partial earnings cases are included in BQC data that an extremely long period would be required to obtain enough partial earnings cases to conduct a useful analysis.

Choosing the Best Procedures - Having identified several feasible procedures, the choice as to the best procedure to use generally must be made on the basis of potential biases and costs. The issues surrounding biases and generalizability are considered below, but note that costs cannot be completely determined until the data requirements of each approach are determined in the next chapter. But that analysis reveals that the proposed approach (the field experiment) is, in fact, the "best" one for this study.

4. **Checks for Biases and Confounding**. The field experiment provides the greatest protection against misleading results. The use of randomization and a control group provides protection against all of the sources of bias identified in this chapter. The drawback of field experiments is that they generally are the most expensive and time consuming of the feasible methods. Interestingly, however, the field experiment in this example proves to be superior to other alternatives, as discussed below.

The quasi-experiment time series approach differs from the field experiment in this example by the absence of a control group and the introduction of additional observation intervals. The additional time is necessary to isolate the effects of external factors. Also, this quasi-experiment requires educating some of the same claimants and employers about two different reporting procedures during the course of the study. These considerations make the time series quasi-experiment more difficult, expensive and time consuming than the field experiment. Thus, the time series approach need not be considered further. (Notice that the intention of a quasi-experiment is to reduce costs. If the full costs of a quasi-experiment equal or exceed those of a true experiment, the latter always should be selected).

The field study using agency data may be subject to several of the biases identified in this chapter. Two major assumptions are required for the results of the field study to be meaningful.

First, it must be assumed that the new procedure will correct the errors that occurred. Second, it must be assumed that the error rates that occurred in the past will occur in the future. Whether these assumptions appear reasonable must be determined by the study team. But they appear to be such strong assumptions that this approach is not used for the TAG example.

5. **Generalizability**. There usually is little interest in study group results alone. The real interest in most cases is not whether some change affected the study group itself, but whether the effects observed for the study group have important policy implications. This means that it is crucial to identify exactly how study results for the proposed field experiment either apply or do not apply more generally. In the example considered, whether the change in partial earnings reporting procedures affects the study group itself is of interest only to the extent that those results can be generalized to a broader group than the study group. In fact, if the results apply only to the (relatively small) study group and not to all partial earnings claims (under reasonable assumptions), or at least an important subset of partial earnings claims, the results would have little relevance for policy purposes. In short, it is important to carefully identify the broader group to which study results may be generalized and to specify the conditions under which the results either do or may apply.

Factors to Consider - The specific factors involved in generalizing study results include carefully defining the larger group or population to which the results apply. It also is important to discuss the assumptions required for the results to apply to even broader groups than the specific study population. For example, study group results generalize specifically to the study population -- all partial earnings claims from which the test and control groups were drawn during the study period. But policy makers are interested in whether the results also are likely to apply to future partial earnings claims, which obviously are not included in the study population. However, assuming the claimants and employers in the study are typical or representative of those likely to be involved in future partial earnings claims, the study results can be used for policy purposes. This means it is crucial that the study sample be carefully drawn to represent the population of interest.

It also means it is important to either be sure that unusual external factors (validity threats) do not affect study results or otherwise to control for those factors in the study. Similarly, the study "treatment" -- the new partial earnings reporting procedure -- must be designed so that it can be routinely used in order to generalize study results. This means that exotic or expensive treatment procedures usually would have little value for a PI study, simply because such procedures would not be operationally feasible. This sometimes means that "exactly correct" and scientifically valid treatment procedures have to be modified in some way to account for the reality of available resources for operational procedures. However, it still is important to use a study approach and procedures that can be replicated in a subsequent study. Thus, ad hoc modifications in the study approach generally should not be made, and such modifications never should be made without carefully documenting them.

6. Selection of the Study Approach. At this point the field experiment appears to be the viable study approach. But the choice actually requires additional information concerning the data requirements of that approach vs. the alternatives. These considerations are discussed in Chapter V.

FOOTNOTES

[1] Elizabethann O'Sullivan and Gary R. Rassel, Research Methods for Public Administrators, New York: Longman, 1989). p. 43.

[2] Ibid. pp. 47-48.

[3] F.N. Kerlinger, Foundations of Behavioral Research. 3rd Edition. New York: Holt, Rinehart and Winston, 1986. p. 373.

BLANK PAGE

CHAPTER V

STUDY PREPARATION

Checklist

1. What data should be collected?

    a. ____ Are the study objectives measurable?

    b. ____ Does each data item relate to at least one study objective?

    c. ____ Are the sampled and target populations the same?

    d. ____ Have any relevant variables been omitted?

    e. ____ Are external factors likely to affect study results correctly accounted for or controlled?

    f. ____ What steps have been taken to reduce external sources of variation that can affect study results?

    g. ____ Will matching or statistical control be used to reduce extraneous sources of variation?

    h. ____ Does the study group correctly reflect the larger group of interest?

2. How much data should be collected?

    a. ____ Is the study group large enough to measure anticipated study effects accurately enough for policy purposes?

    b. ____ Are the expected results reliable? That is, would the same, or nearly the same, results be obtained if the study were repeated?

3. How should the data be collected?

    a. ____ Have both intrusive and nonintrusive procedures been considered?

    b. ____ Have in-person, mail, and telephone approaches been considered?

c. ____ Has the selected approach been clearly described?

d. ____ Have the advantages and disadvantages of he selected approach been discussed?

4. When can the study be completed?

a. ____ Have the critical activities been identified?

b. ____ Has a schedule of study activities been developed using a Gantt chart or some other device?

c. ____ Has a cost and a time been associated with each activity?

5. Does the study still appear feasible?

a. ____ Are the required resources available?

b. ____ Are the required training, staffing, and management support likely to be forthcoming?

CHAPTER V

STUDY PREPARATION

__Introduction__. Several steps are required to successfully plan, conduct and complete a PI study. This chapter deals with the following interrelated issues: what data to collect; how to collect the data; how much data to collect; study completion issues; and study feasibility considerations.

Before we consider the above four issues, it is useful to review the many data variables that need to be identified in order to determine which data to collect for a given study.

1. __What Type of Data to Collect__? In Chapter III, it was recommended that important study variables be identified and diagramed as found in Table III-1. At this time, it is important to classify these study variables in a manner which describes their potential impact on study variation.

While many similar classification schemes can be found in social research literature, the TAG authors feel the following represents the most commonly used terms and classifications. The value of classifying study variables will become apparent as this section of the TAG is reviewed.

__Variables__. The following is a mutually exclusive and exhaustive classification scheme for data items (variables):

a. Explanatory variables
  - cause variables (independent)
  - outcome variables (dependent)

69

b. Extraneous variables
   - controlled variables
   - disturbing variables
   - randomized variables[1]

This classification scheme provides a useful basis for determining what data should be collected in a PI study. Each of these variable categories is discussed below.

Explanatory Variables - The explanatory variables are the focus of any research design. The explanatory variables are related to the goals of the research, and are chosen based on scientific theories or informed knowledge of the problem. That is, they arise from knowledge or insights about the field under study. Explanatory variables sometimes are called experimental variables, and they can be grouped into two distinct sets. The cause variables or the variables that cause the relationship under investigation are the first set (these are often referred to as independent variables). The other set consists of the effect or outcome variables (often referred to as dependent variables) that are affected by the Cause variables in the relationship under investigation.

In the TAG example, the outcome variable is the percent of BYE cases with payment errors (either underpayments or overpayments). The cause variable is the procedure used to obtain the completed BYE form. In general, there may be several cause and outcome variables, but there is only one of each in this example.

Extraneous Variables - Unfortunately, the cause variables rarely, if ever, are the only variables that impact the outcome variable. The other variables that can affect the outcome variable are called extraneous sources of variation, and methods must be devised for separating their effects from the effects of the cause variables. In this example, the following are potential extraneous variables, local office claims load, employer size and industry, employer location, and whether an outside agency manages the employer's UI account. Many other factors also could affect the outcome variable (percent of BYE cases with errors). In an actual study, an attempt should be made to identify and control all such factors that may fall in the extraneous variable category. These factors are likely to be different within each SESA.

Controlled Variables - These are the extraneous variables whose effects the investigator tries to remove by control during the course of the experiment or in the statistical analysis. Devices for control fall under three headings:

1. Constancy (keeping all extraneous variables constant).

2. Outcome matching (matching the groups on all extraneous variables).

3. Statistical control (controlling for group differences in the extraneous variables during the statistical analysis).

Constancy, also referred to as experimental isolation, would require that all extraneous variables be held constant for claimants in both groups (i.e. the claims load could not vary in the local offices, the claimants would all have to work for the same size employer in the same type of industry, etc.). Clearly, constancy is not possible for the TAG example and usually not possible for other PI studies. In any case, constancy probably would be an undesirable control device, since inferences to the broader population of all claimants then would be suspect.

Because of the large number of extraneous variables and the large number of categories for some of the extraneous variables (multiple size categories, multiple industry categories), outcome matching also is not feasible for the TAG study example. Despite the fact that there are a very large number of BYE cases, it would be extremely difficult to match each case involving an earnings report submitted by an employer with an otherwise similar case involving an earnings report submitted by the claimant. Moreover, even if matching were feasible, statistical control (primarily analysis of covariance) still will be more effective as long as a "linear model" is appropriate.[2] Thus, statistical control is the only feasible control device for this example, and it will be discussed further in Chapter VIII.

Deciding exactly which variables should be controlled requires informed judgement by persons familiar with the relationships under study. It obviously is important to remove all major extraneous variables that might exert important effects on the outcome variable(s). In this case, examples of controlled variables that might affect the outcome variable (erroneous BYE payments) could include: the size of the employer; the industry of the employer; the location (urban vs. rural) of the employer; the claims load of the local office; whether the employer's UI account is managed by an outside agency; and the educational level of the claimant. The effects of these extraneous variables can be accounted for by including them as independent variables in the statistical analyses that are conducted (see Chapter VIII).

Randomized Variables - The effects of randomized extraneous variables are handled by randomization and replication. Randomization does not attempt to remove the effect of a source

72

of variation, but instead makes this source act like a random variable that is equally likely to affect either group. A consequence of using randomization rather than control is that larger sample sizes are needed. Hence, randomization generally is used only for those variables that are not major contributors to the variation of the outcome variable.

In this example, claimant characteristics such as occupation and age, although possibly related to the probability that a BYE payment error occurs, probably are not major contributors to whether a payment error occurs. Hence, the effects of these variables might not be explicitly modeled (but instead be left in the model's error term) by assuming they are randomized variables. Hopefully, biases resulting from group differences in these variables would not occur because claimants were randomly assigned to the test and control groups. Then, the only consequence of interest for the problem at hand should be that less variation in the outcome variable can be explained in the study analysis. (This means that a larger sample size is needed to obtain any desired level of statistical precision in the analysis, as discussed in Chapter VIII).

Disturbing Variables - The effects of these extraneous variables are neither controlled nor randomized. In an "ideal" experiment, there are no disturbing variables because they may give misleading results for which tests of statistical significance provide no protection (and no warning), if they are related to (confounded with) the treatment difference. For this example, the most worrisome disturbing variable probably is that the test procedure is a new one so it may involve more errors than would result from implementing the test procedure on a permanent basis. In any case, the possible effects of any disturbing variables should be addressed in the final report.

Summary of Variables - The variables discussed for this example are summarized in Table V-1, on page V-20. It is useful to develop a similar summary table to clarify the relationships among the variables involved in any PI study. The reason is that a careful consideration of the relationship(s) under study usually reveals a number of potential influences that might be missed without a detailed analysis of what factors may affect study results. The example used appears to be a simple one at a general level, but further reflection reveals that (at least) all of the variables in Table V-1 need to be considered in designing and conducting the example PI study. Information on the explanatory and controlled variables must be collected, since they are required for the analysis. Information on the Randomized and Disturbing variables is not absolutely essential. However, it is good practice to collect data on the randomized and disturbing variables whenever possible because that makes it possible to include appropriate qualifications in discussing the results in the final report.

2. How Much Data to Collect? The following three pieces of information generally must be known in order to determine the appropriate sample size for any PI study: an appropriate "level of confidence"; the "variance" of the outcome variable; and an appropriate level of "precision" for the study estimates.

Minimum Sample Size - Unfortunately, it is impossible to provide any simple summary table that defines the minimum or acceptable sample size for various types of PI studies. This is the case because the three items that determine sample size can vary so widely from one study to the next. In fact, selecting the correct sample size is a complex process that requires technical expertise beyond that available in many QC units. However, the required expertise may be available in other units within many

SESAs. Alternatively, the required expertise is readily available in universities, other institutions of higher learning and private sector firms. Because of the complexity involved in identifying the correct sample size for a PI study, the above factors that affect sample size are discussed only briefly in this chapter. But there are many college textbooks that fully develop the issues outlined below (see Bibliography).

The only short cut to having a person with the required expertise determine the correct sample size for any particular PI study is to adopt a safe rule of thumb. This rule of thumb is that a sample of 2,000 is about the largest sample required for almost any study. In fact, a sample size of 400 would be sufficient for many PI studies, and even 400 cases is a larger sample than needed for some PI studies. Thus, analyzing the three elements that determine the appropriate sample size often would reveal that a smaller sample could be used. Accordingly, the issues involved in such an analysis are outlined briefly below.

Level of Confidence - The appropriate level of confidence usually is determined in relation to the budget. Although it theoretically is possible to determine the optimum (cost minimizing) level of confidence, it seldom is done in practice. The basic problem is the extreme difficulty of quantifying the costs associated with making an incorrect decision because the sample was too small. This cost must be quantified to use decision analysis for determining the optimum level of confidence. In practice, it is much simpler to determine the sample size for several alternative levels of confidence. The budgetary implications of each alternative then can be examined.

Variance - Generally, the variance of the outcome variable is easily estimated by statisticians. Past studies, a pilot study, or even a guesstimate concerning the range can be used to develop a reasonable estimate of the variance associated with the outcome variable(s).

Level of Precision - The most difficult item that must be addressed in determining the minimum sample size for a PI study is the appropriate level of precision for study estimates. To determine an appropriate level of precision, some knowledge about the magnitude of the change or characteristic being estimated is required. For this project, a reduction in the error rate for BYE payments from the current level of 32 percent to even 16 percent would be considered a very great success. This means that the sample size must be at least adequate to ensure that the proposed procedure will be identified as the preferred procedure if it actually cuts the error rate in half. If the sample size is not large enough to ensure this, then it is not sensible to even conduct the experiment. This illustrates why it is so important to determine the required sample size in designing and conducting PI studies. Also notice that determining the correct sample size always is relevant unless the entire population is studied, and this rarely is possible in a PI study.

Desirable vs. Optimum Sample Size - A desirable, as opposed to a minimum, sample size can be determined in one of two ways. The first is to determine the sample size required to ensure that the proposed procedure will be identified as the preferred procedure if the proposed procedure reduces the error rate by some amount, say five (5) percentage points. The argument for using this approach is that there is no point in changing procedures unless the change reduces the error rate by at least five (5) percentage points.

The second approach for determining the optimum sample size is to ask how accurate the estimate of the difference between the error rates for the two BYE reporting procedures must be. For example, the sample size may need to be large enough to ensure that the estimate of the difference in the two error rates is off by no more than one percentage point.

Again it is possible to use decision analysis to determine the optimum level of precision. In practice, however, the information required to do this is nearly impossible or, at least, prohibitively expensive to obtain. Practitioners generally determine the sample sizes required for several levels of precision and then examine the budgetary implications of each.

Selecting the Sample Size for the Partial Earnings Example - Table V-2 on page V-21 can be developed to determine the sample size for the BYE example. (See Scheaffer, Mendenhall, and Ott, 1986, for a detailed discussion of how to determine the sample size given the desired level of confidence, the desired level of precision, and the variance of the outcome variable). Table V-2, makes it possible to examine the budgetary implications of each precision-confidence level pair.

In addition, a more complete standard table for determining adequate same size can be found in Appendix C, page 13. However, this table has definite limitations as identified in the Table's cautionary notes.

3. **How Should the Data be Collected**. The question of how data should be collected can be broken into two subquestions. What are the most appropriate data collection methods? What are the most appropriate data collection instruments? Each question is considered below.

Data Collection Methods - Data collection methods can be categorized as follows: (1) personal interviews; (2) telephone interviews; (3) self-administered (usually mail) interviews; or 4) direct observation.[3] The definitions of personal, telephone, and mail interviews are straightforward and require no further explanation. Direct observation, however, requires additional explanation. A good example of direct observation is when a researcher examines hospital records for the incidence of some disease, as opposed to surveying individuals. In instances such as this, the information can be obtained more simply and more accurately by directly observing hospital records. Direct observation also is used in almost all field experiments.

Method Advantages and Disadvantages - Each of the four data collection methods noted above has some advantages and disadvantages, as shown in Table V-3 on page V-22. The choice of the best data collection method is, therefore, dependent on the particular situation being studied. Three general observations can, however, be made.

First, respondents tend to reply more truthfully in situations when their replies are confidential and anonymous. Few persons feel comfortable divulging information the use and exposure of which is unknown. Providing an environment and instructions which give some semblance of anonymity, may increase truthfulness of responses.

Second, mail questionnaires can achieve good results, but the response rates for this method of data collection generally are so low that all reported results are suspect. If a self-administered mail interview is selected as the method for collecting the data, then follow up methods must be devised to ensure an adequate response rate.

Finally, it should be noted that more than one method typically is used to collect survey research data. Mail surveys with personal or telephone follow-up, mail distribution with personal interviews, direct observation with mail or personal follow-up, and many other combinations are possible.

Regardless of the method chosen to administer a PI study questionnaire, it is very important to give adequate consideration to the questions raised in Table V-4 on pages V-23 and 24 when preparing any type of questionnaire.

Data Collection Instrument - Questionnaire design is a very complicated task when more than a few questions or concepts are involved. Two excellent texts on this topic are Schuman and Presser (1981) and Lansing and Morgan (1971). The major concerns in questionnaire construction can be categorized as content, question wording, response form, and question ordering. Table V-4 provides some useful questions to consider with respect to each of these concerns. A few comments on the importance of question wording and sequencing also are provided below.

Consider, for example the following two sets of alternative question wordings:

1a. Do you favor capital punishment?

1b. Do you favor or oppose capital punishment?

2a. If there is a union at a particular company or business, do you think that all the workers there should be required to be union members?

2b. If there is a union at a particular company or business, do you think that all workers there should be required to be union members, or should it be left to the individual to decide whether or not she wants to be in the union?

In general, questions should be balanced and they should not have strong arguments and counter arguments woven into them. Thus, 1b. is preferred to 1a. and 2a. is preferred to 2b.

Question ordering also is crucial. Consider the two following sets of questions:

1a. Will you support an increase in State taxes for roads?

1b. Will you support an increase in State taxes?

2a. Do you feel American Journalists should be allowed free reign in the Soviet Union to report stories back to the United States?

2b. Do you feel Soviet reporters should be allowed free reign in the United States to report stories back to the Soviet Union?[4]

Reversing the order of each set of questions would have an important impact on survey responses.

4. **Schedule of Activities**. The Gantt chart is quite useful in analyzing the feasibility of a PI study schedule. Its advantages

stem from the chart's graphical representation of required study activities, with the length of each bar representing the time required to accomplish the activity. The terms involved in developing a Gantt chart for the BYE study used as the example in this TAG are provided in Table V-5 on page V-25. An example of a Gantt chart for the BYE study is shown in Chart V-1 on pages V-28 and 29. Note that these tables and the Gantt chart include only some of the major activities and resources that would be involved in an actual PI study. Their purpose is simply to illustrate the general concept of developing a Gantt chart. In any case, experience shows that Gantt charts can be valuable tools in keeping studies on track some. Benefits of the Gantt chart include:

a. All jobs are graphically displayed in one easily understood chart.

b. By having the workers or supervisors shade in the percentage of the bar that corresponds to the percentage of the activity completed each work period, it is easy to check the overall progress of the system of activities at any point. (Note that, for display purposes, the bars in Chart V-1 are fully shaded. None would be shaded at the beginning of a study).

c. When the required resources are limited, the chart provides an initial evaluation of the planned use of the resources.

The project evaluation and review technique (PERT) extends the Gantt chart by determining the probability that a project can be successfully completed within a particular time frame. For details on PERT see Richmond, 1968, chapter 17. PERT is not

BLANK PAGE

discussed in this TAG, but it can be useful in planning and conducting PI studies in some circumstances.

Study Feasibility Revisited - The Gantt chart and related Tables V-5, V-6 and V-7 clearly lay out the activities, resources, and costs to complete the field experiment. Notice that obtaining the information required to complete these tables and the Gantt chart involves considerable work for a study of any magnitude. But the resulting information is required to assess study feasibility in terms of resource requirements, study completion time and total study costs. This allows management to make an informed decision about whether the study should be conducted.

5. **Study Management and Control**. The use of a bar chart, such as a Gantt chart, is critical beyond the resource planning and time allocation phase. Such a chart is invaluable in managing and controlling a study to its conclusion.

As specified in paragraph 4.b. under Schedule of Activities, the Gantt chart is best used as a management tool by having the empty bars shaded as work on each task progresses. As a result, the study manager can at a glance, determine the status of each task in the study and take corrective action if necessary to get the study back on schedule.

A very important aspect of successfully using the Gantt chart to its full potential is gathering the information necessary to use it in a timely fashion. In order to do that, the project manager must be aware of the need for open and regular communication with study principals. Without effective study communication channels, the Gantt chart as a study management and control tool loses its effectiveness.

82

With respect to the Gantt chart, the study manager's primary responsibility is to identify problems in the following two areas:

a. Scheduling of study tasks -- is the study on target?

b. Resource allocation -- are resources sufficient to stay on schedule?

When a variance is detected, the study manager must immediately work to develop possible alternatives to avoid interfering with the study completion date.

6. Threats to a PI Study - Potential problems or threats to a PI study after it begins can be grouped into two areas, internal and external. External threats are potential problems that are outside of the control of the study manager. Examples are other unexpected priorities, workload increases, staffing changes, or a shift in agency priorities.

Internal threats, on the other hand, are potential problems that are at least to some degree within the control of the study manager. The study manager must anticipate and attempt to prevent these problems from occurring.

a. **Underestimation of time to complete tasks** - Care must be taken to avoid underestimating the time it will take to complete study tasks. Underestimation can create stress and frustration as task completion dates arrive with tasks unfinished. Underestimation of time to complete tasks can be avoided by carefully considering staff's time estimates in the resource allocation and Gantt charting phase of study development.

b. **Organizational Politics** - The study manager must be fully aware of the primary and secondary staff necessary to involve in the study in the development and maintenance of organizational support. However, the study manager must also consider the informal organization inherent in any organization. The informal organization reflects the informal relationships that develop between people in an organization to meet the needs of the organization's members, such as companionship, belonging and other social needs[6].

The goal of the study manager is to identify and involve all formal and informal organization members who have an interest in the study. The purpose is to develop their support and thereby avoid possible resistance later in the study.

c. **Resistance to Change** - Since organizational members are likely to perceive a study as potential change, they may view a study as a threat to their jobs and/or the existing system. Study managers can go a long way in defusing the resistance to change by:

(1) Being sure organization members understand the nature and purpose of the study through open communications.

(2) Allowing for staff participation in study development in order to foster organizational support for the study.

(3) Stressing the potential benefits of the study to organization members. Employees who care about the quality of the UI system will be more likely to support the study if they understand how it may benefit the system and the people they serve.

d. **Employee Morale and Motivation** - In order to build study support and a sense of study ownership by study participants, the study manager is encouraged to be participatory in his/her approach and involve study staff as often as possible in every phase of the study. This also contributes to the development and maintenance of organizational support through the study.

85

TABLE V-1

SUMMARY OF STUDY VARIABLES FOR BYE EXAMPLE

A. Explanatory Variables

Outcome variable (dependent)

Percent of partial claims with payment errors

Cause variable (independent)

Normal and test procedures for reporting partial earnings.

B. Extraneous Variables

Controlled Variables

Employer size
Employer industry
Employer location
Local office claims load
Whether outside agency manages UI account for employer
Claimant education

Randomized Variables

Claimant age
Claimant occupation

Disturbing Variable

The procedure is new.

TABLE V-2

ALTERNATIVE SAMPLE SIZES FOR THE BYE REPORTING PROCEDURES STUDY

		CONFIDENCE LEVEL			
		80%	90%	95%	99%
PRECISION	± 1%	4,225	6,889	9,900	16,576
	± 3%	469	765	1,089	1,842
	± 5%	169	275	384	663
	± 10%	42	67	96	166

TABLE V-3

ADVANTAGES AND DISADVANTAGES OF ALTERNATIVE DATA COLLECTION METHODS

A. Personal Interview

Advantages
1. People are more apt to respond when confronted in person.
2. The interviewer can note specific reactions and eliminate misunderstandings.

Disadvantages
1. High time and dollar cost.
2. The interviewer may introduce bias.
3. The interviewer may record answers incorrectly.

B. Telephone Interview

Advantages
1. Low time and dollar cost.
2. Interviewer can be monitored to be sure protocol is followed.

Disadvantages
1. Difficult to find a sampling frame.
2. Interview must be kept short.
3. It may be difficult to eliminate misunderstandings.

C. Self-Administered (Mail) Interview

Advantages
1. Low time and dollar cost.

Disadvantages
1. Low response rates.
2. Questions must be very clear.
3. It may be difficult to eliminate misunderstandings.

D. Direct Observation

Advantages
1. Accurate data.

Disadvantages
1. High time and dollar cost.
2. Not always possible.

TABLE V-4

GUIDE FOR QUESTIONNAIRE CONSTRUCTION*

A. Content

1. Is each question necessary? How will it be useful?
2. Are several questions needed on the subject matter of each question?
3. Do respondents have the information necessary to answer the question?
4. Does the question need to be more concrete, specific, and closely related to the respondent's personal experience?
5. Is the question content sufficiently general and free from spurious concreteness and clarity?
6. Do the replies express general attitudes and only seem to be as specific as they sound?
7. Is the question content biased or loaded in one direction, without accompanying questions to balance the emphasis?
8. Will the respondents give the information that is asked for?

B. Wording

1. Can the question be misunderstood? Does it contain difficult or unclear phraseology?
2. Does the question adequately express the alternatives with respect to the point?
3. Is the question misleading because of unstated assumptions or unseen implications?
4. Is the frame of reference clear and uniform for all respondents?
5. Is the wording biased? Is it emotionally loaded or slanted toward a particular kind of answer?
6. Is the question wording likely to be objectionable to the respondent?
7. Would a more (or less) personalized wording of the question produce better results?
8. Can the question be better asked in a more direct or a more indirect form?

C. Form of Response

1. Can the question best be asked in a form calling for check answer, short (word or two) answer, free answer, or check answer with follow-up free answer?
2. If a check answer is used, which is the best type for this question -- dichotomous, multiple choice, or scale?
3. If a check list is used, does it cover adequately all the significant alternatives without overlapping and in a defensible order? Is it of reasonable length? Is the wording of the items impartial and balanced?
4. Is the form of the response easy, definite, uniform, and adequate for the purpose?

D. Sequence of Questions

1. Is the answer to the question likely to be influenced by the content of the preceding questions?
2. Is the question led up to in a natural way? Is it in correct psychological order?
3. Does the question come to early (or too late) from the point of view of arousing interest and receiving sufficient attention, avoiding resistance, etc.?

*Based on Selltiz, et al., Research Methods in Social Relations, pp. 552-574.

TABLE V-5
MAJOR ACTIVITIES FOR THE BYE PI STUDY

ACRONYM	DESCRIPTION	DURATION	COST (PERSON DAYS)
DEVFORM	Design employer form	1 wk	5
TESTFORM	Field test employer form	1 wk	5
SELEMP	Select sample of employers	12 wks	15
DISTFORM	Distribute forms to employers	2 wks	10
DEVPROC	Develop written procedures for local offices	2 wks	10
TRNSTAFF	Train local office staff	2 wks	10
TESTPROC	Field test local office procedures	13 wks	5
COLDATA	Collect data in local offices	13 wks	10
INVESCLM	Investigate selected claims	13 wks	85
CODEDATA	Encode data	3 wks	15
ANALDATA	Analyze data	4 wks	20
FNLREPT	Prepare final report	4 wks	20

TABLE V-6
PERSONNEL RESOURCES AVAILABLE

ACRONYM	DESCRIPTION	AVAILABILITY (Person Days)
FCSPEC	Forms Control Specialist	1
SYSANAL	Systems Analyst	15
RESANAL	Research Analyst	72
DEOPR	Data Entry Operator	15
INVES	Investigator	85
LODEP	Local Office Deputy	22

BLANK PAGE

TABLE V-7
ESTIMATED RESOURCE USAGE FOR MAJOR ACTIVITIES

JOB (ACTIVITY)	RESOURCE PERSON DAYS REQUIRED FOR ACTIVITY					
	FCSPEC Qty	SYSANAL Qty	RESANAL Qty	DEOPR Qty	INVES Qty	LODEP Qty
DEVFORM	1		4			
TESTFORM			5			
SELEMP		15				
DISTFORM			5		5	
DEVPROC			5			5
TRNSTAFF			5			5
TESTPROC			3			2
COLDATA						10
INVESCLM					80	
CODEDATA			5	15		
ANALDATA			20			
FNLREPT			20			

CHART V-1
GANTT CHART FOR MAJOR ACTIVITIES FOR EXAMPLE PI STUDY

JOB (ACTIVITY)	DATE SCHEDULED*																		
		OCTOBER				NOVEMBER					DECEMBER				JANUARY				
		5	12	19	26	2	9	16	23	30	6	13	20	27	4	11	18	25	
DEVFORM		[-]																	
TESTFORM				[--	---]														
SELECTEMP		[----	----	----	----	----	----	----	----	----	----	----	---]						
DISTFORM														[-]					
DEVPROC		[---	--]																
TRNSTAFF				[--	---]														
TESTPROC						[-]													
INVESCLM							[---	----	----	----	----	----	----	----	----	----	----	----	-->
CODEDATA																[---	----	----	-->
ANALDATA																			-->
FNLREPT																			-->

CHART V-1 - PAGE 2
GANTT CHART FOR MAJOR ACTIVITIES

JOB (ACTIVITY)	DATE SCHEDULED*																	
	FEBRUARY					MARCH												
		5	12	19	26	2	9	16	23	30	6	13	20	27	4	11	18	25
DEVFORM																		
TESTFORM																		
SELECTEMP																		
DISTFORM																		
DEVPROC																		
TRNSTAFF																		
TESTPROC																		
INVESCLM		--]																
CODEDATA		[-----]																
ANALDATA				[----------]														
FNLREPT							[---------]											

FOOTNOTES

[1] W.G. Cochran, Planning and Analysis of Observational Studies. New York: John Wiley & Sons, 1983, p. 69.

[2] Ibid. Chapter VI.

[3] R.L. Scheaffer, W. Mendenhall and L. Ott, Elementary Survey Sampling. 3rd Edition. Boston: Duxbury Press, 1986, pp.27-29.

[4] Ibid. p. 33.

[5] D.D. Bedworth, Industrial Systems: Planning, Analysis, Control. New York: The Ronald Press Company, 1973, p. 37.

[6] A.J. DuBrin, R.D. Ireland, J.C. Williams, Management and Organization, 6th Edition. Cincinnati: South-Western, 1989, p. 24.

CHAPTER VI

PROPOSAL PREPARATION

Checklist

1. Is the study problem fully explained?

   a. ____ How did QC data analysis help identify or confirm the problem?

   b. ____ How has analysis of other UI data helped identify or confirm the problem?

   c. ____ How will a PI study help address the problem?

   d. ____ Why is additional study necessary before a program improvement action can be taken?

2. Are study goals and objectives clearly defined?

   a. ____ Are study objectives explained?

   b. ____ How will the achievement of study objectives be measured?

   c. ____ How will achieving study objectives address the problem?

3. Is the study approach clearly explained?

   a. ____ How will the study approach provide the necessary information?

   b. ____ Are data sources identified?

   c. ____ Is sampling methodology explained?

   d. ____ Have all data collection and recording procedures been explained?

   e. ____ Are data analysis plans explained?

97

4. Is a detailed resource estimate and corresponding work plan prepared?

   a. ____ Is a study start date and an expected completion date specified?

   b. ____ Are staff requirements estimated?

   c. ____ Are other resource requirements estimated?

   d. ____ Is the size of the reduction in QC sampling explained?

   e. ____ Have you demonstrated how you arrived at your resource estimates?

   f. ____ Is the work plan flexible enough to accommodate unforeseen events?

   g. ____ Does the work plan establish interim goals and benchmarks to keep the study on schedule?

   h. ____ Does the work plan specify how the results of the study will be presented to agency administrators?

CHAPTER VI

PROPOSAL PREPARATION

Program improvement study proposal requirements and the content outline for such proposals are contained in the current UI program letter issued by the UIS National Office (Appendix A). The proposal areas covered and the discussion in this chapter are intended to be illustrative and not a formal interpretation of those UIPL guidelines. The proposal sections listed below will vary in length and detail depending upon the complexity of the study. But adequately addressing the applicable items on the checklist for this chapter will expedite the PI review/approval process.

1. Introduction. An adequate introduction briefly summarizes the proposed PI study. Although brief, the introduction is the key to attracting attention to the value of the project. Avoid jargon and cover the highlights of the proposal. The body of the proposal should follow as closely as possible the required content described in the appropriate UIPL.

2. Problem Identification. This section should explain the process used to define the problem and develop the study objectives. The narrative should address how the analysis of QC data contributed to this process and what, if any, auxiliary data sources provided supporting documentation of the problem. Indicate the specific problem or UI program facet that is proposed for the PI study. Indicate why further study is needed before SESA management can consider potential PI actions. Well designed tables, charts and graphs can be useful in highlighting important details surrounding the problem area. Raw data and

involved data presentations should be placed in appendixes for review by interested individuals.

3. **Study Objectives and Methodology**. A number of research designs can be employed for a PI study. Discussion of how the methodology selected will achieve the broad goals and specific objectives of the study is essential. Additionally, the target population and sampling routine are major study components that must be described in the proposal. What are the study objectives? What does the SESA expect to learn from the study? Describe the methodology to be followed. Explain how this study approach will provide additional information needed by UI management either to define a specific problem/issue or to take a program action. Data analysis techniques are an integral part of the methodology that will be utilized at the conclusion of the assessment period. Presentation of these techniques should be sufficiently detailed to provide the reviewer with a basic understanding of the methods that will be used to evaluate the study findings.

4. **Administrative Considerations**. The final section of the PI proposal will communicate three important areas of study development: (1) the anticipated results of the study, including how results will be presented to SESA administrators; (2) a detailed estimation of the total resources that will be dedicated to the study, with a distinction made between the resources that will come from BQC and those that will be committed from regular SESA staffing; and (3) a work plan that specifies the study start date, specific time frames for completing major tasks and the anticipated date of study completion. The following excerpts from the current appropriate UIPL specify the minimum content for administrative sections of a PI proposal:

a. **Results Expected** - What results are anticipated? How will the results of the study be presented to UI principals? How does the SESA anticipate translating this information into decisions that may lead to UI program improvement?

b. **Projected Study Resources** - How many QC staff and other UI resources will be needed to conduct the program improvement study? Be sure to identify QC positions the SESA plans to dedicate to the study and indicate the size (number of cases) of the reduction in QC weekly sampling requested during the study period.

c. **Work Plan** - What is the projected time period for implementation of the PI study and preparation of a final study report? Present a study work plan that specifies QC staff hours and time (in weeks) required for completing major tasks or segments of the study and QC staff time, if any, needed for preparation of the final report. Indicate the anticipated study start date, study completion date, and planned date for completing the final report.

TABLE VI-1

MINIMUM PI PROPOSAL REQUIREMENTS

The minimum requirements for a PI proposal are specified in UIPL 37-91. Because these minimum requirements must be met to obtain approval for any PI study, they are reported in outline form below.

MINIMUM REQUIREMENTS FOR A PI PROPOSAL

A. **Problem Identification**
   (1) Statement of problem
   (2) BQC data pointing to the problem
   (3) Other pertinent UI data analyses

B. **Objectives and Methodology**
   (1) Study objectives
   (2) Expected insights from the study
   (3) Methodology/study approach
   (4) Potential value of study for UI management
   (5) Explain/justify study approach

C. **Potential Results**
   (1) Results expected
   (2) Presentation of findings to management
   (3) Likely program improvement actions

D. **Study Resources**
   (1) Staff requirements
   (2) QC sample reduction needed

E. **Work Plan**
   (1) Schedule of key activities
   (2) Deployment of QC staff
   (3) Study start and completion dates
   (4) Final report date

CHAPTER VII

DATA COLLECTION AND RECORDING

Checklist

When collecting and recording data, the following issues must be considered.

1. Were any inaccuracies introduced because of the study design?

   a. ____ Has the realism of the experiment been considered?

   b. ____ Has anything been done to eliminate nonresponse biases?

   c. ____ Are the anticipated results likely to accurately reflect the actual situation (e.g., will respondents truthfully respond)?

2. Were any inaccuracies introduced during the collection process?

   a. ____ Were training materials prepared?

   b. ____ Are data gathering instructions and questions clear?

   c. ____ Were all materials and procedures pretested?

   d. ____ Were positive incentives given to carry out all instructions, and were no incentives given to circumvent instructions?

   e. ____ Was a sample of the data or respondents checked to verify that interviewers conducted the interviews properly?

   f. ____ Was a sample of the data or respondents checked to verify that interviewers recorded their responses accurately?

3. Were any inaccuracies introduced during the encoding process?

   a. ____ Was all keypunching verified?

   b. ____ Was an edit program written?

c. ____ Have procedures been developed to secure and back up the data?

d. ____ Was a sample of the encoded questionnaires verified back to the original questionnaires?

CHAPTER VII

DATA COLLECTION AND RECORDING

Every statistical analysis presumes accurate data. If the data are not accurate, not only is the conclusion incorrect, but gathering more data only serves to make the analyst more confident of the incorrect conclusion. The adage that "the only thing worse than not knowing something, is knowing something that just ain't so" certainly applies to PI studies.

Before the data is analyzed, therefore, it is very important to determine whether any factors are present that might have resulted in inaccurate data. Generally the causes of inaccurate data can be categorized as resulting from: (1) a poorly designed experiment or survey instrument; (2) poorly trained data collectors or interviewers; or (3) poorly designed verification procedures. Each of these possible causes of inaccurate data are examined below.

1. Design Accuracy. The main factors that contribute to inaccurate data at the experimental design stage are the selection of an inappropriate group for the experiment or a lack of realism for the experiment. For example, asking for volunteers to attend a job finders workshop, and then measuring the number of contacts made during the week that the claimants participated in the workshop would result in inaccurate data for at least two reasons: volunteers almost certainly are not typical of all claimants (an inappropriate test group); and claimants typically do not have a trainer available to help them make job contacts (an unrealistic test).

The main factors that contribute to inaccurate data in survey research are: (a) an inappropriate group of respondents; (b) a high rate of nonresponses; (c) poorly worded questions that are not understood; (d) or poorly designed questions that claimants will not answer truthfully (either because the question leads the claimant or because the claimant refuses to answer truthfully).

A "one-shot" survey of claimants who had both earnings and benefits in the same quarter would, for example, suffer from all of the above problems. In particular: (1) claimants who had both earnings and benefits in the same quarter probably are not typical of all claimants; (2) claimants who respond to the first mailing of a survey may differ from claimants who do not respond; (3) some claimants may not understand questions written in English; and (4) claimants who had unreported earnings are unlikely to admit that they did not report all of their earnings.

For the BYE example used in the TAG, it seems likely that accurate data can be collected. Random selection of claimants exposed to the test and control procedures provides protection against the selection of inappropriate groups of entities, and collection of the data in an operational setting assures the realism of the experiment.

2. **Collection Accuracy**. Training manuals may not be required for the TAG example since it is not an example of survey research. Consistency also is not a problem, and incentives are not a problem.

3. **Encoding Accuracy**. This section is as relevant for field studies as it is for survey research. The following issues are relevant for this section:

- Deciding on how blanks, zeros and missing data are coded.

- Verifying all keypunching.

- Checking a sample of the data to verify data recording procedures. (The size of the sample for the example is large enough so that there is only one chance in 100 that 1% or more of the data items have errors).

- Developing a computer edit program would be simple for this problem because very few data items are being collected and they are not contradictory. (One possible edit is that partial earnings cannot be greater than the WBA).

- Detecting any problems associated with getting data in the wrong fields by calculating univariate means and extreme values for the data. (There should be no missing values).

A check on how well procedures are defined is to see if someone else could replicate the procedures. In this example, it seems clear that someone could.

BLANK PAGE

CHAPTER VIII

DATA ANALYSIS

Checklist

1. Is the intention to simply describe what happened with the study group or to make inferences to a larger group of interest?

   a. ____ Is the target population for policy purposes clearly defined?

   b. ____ Is the sampled population clearly defined?

   c. ____ Does the sample correctly represent the target population?

2. For situations in which the intention is to use the study results for generalizing to a larger group of interest.

   a. ____ Have the effects of nonsampling errors (e.g., data gathering errors) been clearly identified and minimized?

   b. ____ Has either matching or statistical control been used to control external factors that can influence study results?

   c. ____ Are confidence intervals constructed for all study estimates?

   d. ____ Could the width of the confidence interval be reduced by using a "linear model" to explain some of the variance in the study results?

3. Are the results of the data gathering and analysis clearly presented?

   a. ____ Are graphs or other visual aids used whenever possible to highlight the findings?

   b. ____ Are all graphs and tables clearly labeled and self-explanatory?

4. Should any action be recommended?

a. ____ Are the results statistically significant?

b. ____ Are the results of any practical importance?

c. ____ Have alternative recommendations been considered?

d. ____ Has the risk or uncertainty associated with each recommended action been clearly identified?

e. ____ Have the costs and benefits associated with each recommended action been clearly identified?

f. ____ Do expected benefits exceed expected costs for any recommended actions?

CHAPTER VIII

DATA ANALYSIS

Data analysis involves several steps that are overviewed in this chapter. Most PI studies involve samples because analyzing the entire population usually would be far too costly and time consuming. Thus, the issues involved in describing the population based on a sample, which is the group of interest for policy purposes, are discussed in the first section. Then, the reliability and validity of PI study estimates are discussed, followed by some tips for effectively presenting study results. Finally, the issues involved in deciding whether to recommend any actions based on study results are discussed.

1. Sampled vs. Target Population. The objective of nearly all data analysis is to describe a specific population. In some instances it is possible to measure every element in the population (take a census). For example, the amount of every UI check written to each claimant is recorded for the entire population.

When a census is taken, the objective of the data analyst is to make the information contained in the set of data stand out for easy comprehension. For example, no one wants to see thousands of pages containing the amounts of the individual UI payments. It is far more informative to see summary measures, such as the average or total amounts paid, or graphs of summary measures, such as bar charts showing average and total amounts paid. Summary measures and graphic displays are the main tools available to data analysts for facilitating comprehension.

Most of the time, however, it is not possible to take a census. In the UI system, it is not possible to audit every payment to determine whether it was paid correctly. Time and cost considerations make it necessary to assess the accuracy of all payments on the basis of a very small portion (sample) of that population, as is done in the BQC program.

Describing a population on the basis of a sample, unfortunately, is somewhat complicated because it is necessary to determine:

(1) the most appropriate summary measures and graphic displays to utilize for the sample data (which also is required when using population data); and

(2) how accurate the sample description is for the population.

A statement about a population that is made on the basis of a sample is called a statistical inference (something is inferred about the population on the basis of the sample). The techniques of statistical inference are based primarily on probability theory, which is used to quantify the risk or uncertainty associated with the inference. For a statistical inference to be complete, four components are necessary:

-- a target population;

-- a sampled population;

-- an inference; and

-- an indicator of reliability.[1]

Target Population - The target population is the population of interest for the PI study. In this example, the target population is all intrastate BYE claims. The only restriction on defining the target population is that its definition must be complete enough so that it is possible to classify any element as either belonging or not belonging to it.

Sampled Population - Statements about a target population generally must be based on sample data. For example, red blood cell counts obviously must be determined from a sample, since the target population (all of a person's blood) cannot be drawn for this purpose. Similarly, it is necessary to use samples for most PI studies, including the BYE example used in the TAG. In this case, there simply are not sufficient resources available to check every BYE claim for errors. In any case, such an approach would waste resources since the information required for policy purposes can be obtained by examining a relatively small sample of all BYE claims.

It is very important that the sampled and target populations be identical or nearly identical. This is the case because direct inferences can be made only to the sampled population. A classic example of an instance where the sampled and the target populations were not the same is the telephone survey used by Reader's Digest to incorrectly conclude that Hoover would defeat Roosevelt in the 1932 presidential election. The target population was all registered voters who would vote in the election. The sampled population was households with telephones. Clearly, the sampled population overrepresented high income households and underrepresented low income households. Equally clear is the fact that calling more households would not have made the survey an accurate predictor. On the contrary, it simply would have made the pollsters more confident of their incorrect prediction.

The target and sampled populations also differ for the BYE example, because the relevant policy issue relates to both current and future BYE claims, but the sample can be drawn only from current BYE claims. Since future BYE claims cannot be sampled, the ever dangerous "ceteris paribus" or "other things equal" assumption must be invoked. If the future is not at least very similar to the present in terms of factors that can affect BYE claims, then study results may not be valid for future BYE claims. Regardless of the future, however, the sampled and the target populations are identical for current BYE claims, and a valid probability sample can be from the population of current BYE claims. Also, it appears reasonable (after some analysis by State experts) to assume the sampled and target populations are likely to be very similar, even if they are not identical.

Inference - Making a statement about the reduction in the error rate for BYE claims that would occur as a result of introducing the proposed change in the filing procedure for partial claims is the primary focus of the example PI study. The study estimates will be based on a sample of current BYE claims, and these estimates will be used to make statements (inferences) about the sampled population of current BYE claims. Then, it will be assumed that these inferences for the sampled population also apply to the target population (of both current and future BYE claims).

Reliability - Although the observed reduction in the BYE error rate for the sample is the best guess for the reduction in that error rate for the population, this guess generally will not be exactly correct. For example, if a fair coin is tossed 1000 times, the best guess for the number of heads is 500, but the probability that exactly 500 heads appear is only 2 out of 100. Because the best guess in this example or any other involving sample data usually is not exactly correct, it is necessary to place bounds on the likely error associated with the best guess (point estimate). In the coin tossing example, the number of heads in 1,000 tosses will be between 470 and 530 for 95 percent of the times the experiment is run. This means that it is possible to be 95 percent "confident" that the best guess (point estimate) will be off by no more than 30 (either above or below 500). Any time a probability sample is selected, it similarly is possible to state how "reliable" the point estimate is -- that is, the confidence level(s) for how much error likely is associated with the point estimate.

The bounds for the error of a point estimate can be used to construct a confidence interval. For example, suppose that, for a sample of BYE claims, 15 percent contain some type of error. Furthermore, assume that it is possible to be 95 percent confident that this point estimate (best guess based on the sample) differs from the true error rate for the population by no more than two percentage points. This means it is possible to be 95 percent confident that the error rate for the population is between 13 percent (= 15% - 2%) and 17 percent (= 15% + 2%). The 13-17 percent interval is called a 95 percent confidence interval in this example.

The necessity of using and understanding confidence intervals whenever sample data are utilized is so important, although often ignored, that this section concludes with the following observation:

> Whenever any statement about a population is based on a sample (a statistical inference is made) a confidence interval MUST be constructed. Note that the ONLY reason why a confidence interval must be constructed is because it is not possible to be certain that the sample estimate is exactly correct for the population. When a census is taken, it is neither necessary nor possible to construct a confidence interval. But confidence intervals ALWAYS should be constructed for estimates based on sample data.

2. **Reliability and Validity**. It is useful to distinguish between the reliability and the validity of statistical inferences. Reliability relates to sampling errors, and validity relates to nonsampling errors that were discussed in Chapter VII. The sampling errors discussed in this chapter almost always are much easier to control and quantify than nonsampling errors. In any case, both types of errors are briefly discussed below to clarify the distinction between validity and reliability.

Validity - Recall that validity relates to nonsampling errors or external threats to validity, as discussed in previous chapters. Briefly, realism in the TAG example was assured by conducting a field experiment, and biases in the groups selected for study were controlled by randomization and stratifying on key factors such as firm size, firm industry, and how the firm's UI account is handled. Univariate checks on the variables during the data analysis provided further assurance that differences in the two groups are not due to selection biases.

Reliability - Reliability can be illustrated by again considering the coin tossing example described above. However, if the coin is fair, 50 percent of the tosses will result in heads after enough tosses. However, if we toss the coin only 5 times, it is possible to get no heads (probability = 0.03) or all heads (probability = 0.03). Thus, any estimate of the "percent of tosses that will result in heads" based on this sample of 5 tosses would not be very reliable. In contrast, if the coin is tossed 1000 times, it is unlikely that fewer than 47 percent heads will occur (probability = 0.025) or more than 53 percent heads will occur (probability = 0.025). Thus, an estimate of the "percent of tosses that will result in heads" based on 1000 tosses would be very reliable.

General Confidence Interval Rules - The above discussion can be generalized by noting that, if the target and sampled populations do not differ and a probability sample is selected, then it always is possible to: (1) place a bound on how far off the sample result is likely to be; and (2) make that bound as small as desired by increasing sample size. For example, a 95 percent confidence interval for the reduction in the BYE payment error rate that would result from the new BYE filing procedure is constructed and reported in Table VIII-1 on page VIII-17. This interval incorporates all sampling error, but ignores all nonsampling error.

Reducing Confidence Intervals by Using the Linear Model - Except for the situation where paired differences are analyzed (or differences are unimportant), the data should be analyzed using a linear model. The reason for this is that the width of the confidence interval for the difference between two population means or proportions depends upon the variances of the populations being compared. Unfortunately, using the linear model requires expertise that may exceed that available in some BQC units, although the SPSS statistical package provided to all States does have the software needed to use the linear model. In any case, the required expertise may be found in other units within the SESA, such as the UI research unit.

The variance of the population is estimated around the regression line with a linear model, rather than around the population mean (which occurs when no model is used). If there are several important causal variables, the variation around the regression line will be much smaller than the variation around the mean, as illustrated in Figure VIII-1 on page VIII-21. This results in much more precise estimates (i.e., narrower confidence intervals) than if a linear model is not used.

For the BYE example, a possible linear model can be specified as:

$$\begin{aligned}&\text{Payment Error} = B_0 + B_1 * \text{Procedure}\\&+B_2 * \text{Size} + B_3 * \text{Industry}\\&+B_4 * \text{Location} + B_5 * \text{Load}\\&+B_6 * \text{Preparer} + B_7 * \text{Education}\\&+B_8 * \text{Age} + B_9 * \text{Occupation}\\&+e\end{aligned}$$

where Procedure, Size, Industry, etc. are the variables included in Table V-1. If these cause and controlled variables are importantly related to the occurrence of BYE payment errors, the confidence interval for B_1 (which measures the impact of using the proposed procedure rather than the existing procedure) will be narrower than the confidence interval for $P_1 - P_2$ (where P_1 measures the percent of BYE cases with errors under the existing procedure and P_2 measures the percent of BYE cases with errors under the proposed procedures). In fact, if the effective range of Y, given X, is only 1/3 the effective range of Y (as shown in Figure VIII-1) then (assuming normality) the width of the confidence interval for B, will be only 1/3 the width of the confidence interval for $P_1 - P_2$. This is of more than academic interest because it has important implications for study costs. In this example, the linear model makes it possible to reduce the confidence interval to 1/3 the size it otherwise would be. Thus, the sample size needed (to achieve the same level of confidence) is much smaller if the linear model is used than if it is not. This, of course, translates into resource and cost savings that are important considerations in any PI study.

3. Presentation of Results. Tabular and graphical formats should be used whenever possible to organize, summarize, and present numerical data. However, it also is important to clearly/accurately label all tables and graphs. The importance is partly indicated by the fact that William S. Cleveland found, in a study in which a large number of graphs were reviewed, that 30 percent of the graphs contained one or more errors in the following areas:

(1) Explanation (15.4%) -- something on the graph is not explained.

(2) Discrimination (10.1%) -- items on the graph, such as different symbol types, can not be easily distinguished due to the design or size of the graph.

(3) Construction (6.4%) -- A mistake was made in the construction of the graph such as tick marks incorrectly spaced, mislabeling, items omitted, and wrong scales.

(4) Degraded image (6.4%) -- some aspect of the graph was missing or partially missing due to poor reproduction.[2]

To safeguard against these problems, Cleveland offers 28 suggestions which can be summarized into the following four rules:

(1) Clear vision -- Use uncluttered graphs with prominent symbols.

(2) Clear understanding -- Make all legends comprehensive and informative.

(3) Clear scales -- Choose appropriate scales

(4) Iterate -- It normally takes several tries to develop a clear graph.[3]

For the TAG example, no graphs or tables are developed, but Table VIII-1 is developed to summarize the point estimates and associated confidence intervals to again emphasize the importance of always reporting confidence intervals for sample estimates.

4. Recommended Actions. Given the results shown in Table VIII-1, two questions naturally arise: (1) are the results of any practical importance? and (2) do the results justify taking any action? Each of these questions is considered below.

Statistical Significance vs. Practical Importance - In general, results are statistically significant (at the given confidence level) whenever the confidence interval does not include 0. For example, if the reported confidence interval for the partial payments example were 0.5 percent (lower limit) to 1.0 percent (upper limit), the results would be statistically significant in that the proposed procedure very likely has some effect larger than zero. However, these particular results probably would not be of any practical **importance** for policy purposes. It is doubtful that any SESA would implement the proposed procedure if it reduced the BYE payment error rate by only 1 percentage point (from 32 percent to 31 percent).

Statistics can be used to isolate the maximum or minimum magnitude of the result, but it cannot be used to assess the practical importance of the result. The practical importance of a result must be assessed on the basis of direct understanding of the problem as it relates to UI operations. In the example illustrated in Table VIII-1, the procedure very likely will reduce the BYE error rate by at least 11 percentage points. Given the magnitude of this reduction and the magnitude of BYE errors, this result certainly would appear to be of some practical importance. Nevertheless, it still may not be reasonable to implement the test procedure. To determine whether action is warranted, the costs and benefits of the proposed procedure must be examined more closely.

Cost-Benefit Analysis - Cost-benefit analysis (CBA) of some sort is required to estimate whether the likely benefits of a potential change are expected to exceed the likely costs. There are numerous texts that discuss the issues involved in applying cost benefit analysis to public programs. A particularly useful statement for applying cost-benefit analysis in PI studies is found in Sassone and Schaffer, 1978:

> Among noneconomists, "cost-benefit analysis" and "cost-effectiveness analysis" are often erroneously considered to be techniques for appraising public projects. If CBA (cost-benefit analysis) is to be considered a technique, it is at best a loosely defined one. A "cost-effectiveness analysis" is considered to be a special form or subset of CBA distinguished by the difficulty with which project benefits can be identified in terms of dollars.
>
> We define cost-benefit analysis as an estimation and evaluation of net benefits associated with alternatives for achieving defined public goals... Cost-benefit analysis is a generic term embracing a wide range of evaluative procedures which lead to a statement assessing costs and benefits relevant to project alternatives. The variety of problems addressed and the ingenuity which must be exercised in estimating costs make it particularly difficult, if not impossible, to design an all-purpose CBA procedure.[4]

The definition of CBA given by Sassone and Schaffer has proven useful to the authors of this TAG, who also agree that projects differ so much in character that an all-encompassing procedure cannot be defined. However several general principles may be stated, and a number of guidelines have been established over the years. What follows is a discussion of those principles and guidelines as they apply to PI studies.

In order to conduct a CBA, it is necessary to quantify the costs and benefits of the program, and amortize them over the expected life of the program. The costs and benefits resulting from implementing a PI change can be classified as either internal or external. Internal benefits and costs accrue directly or indirectly to the entity under study (the state UI agency). External costs and benefits escape the agency and fall into the hands of others. External benefits may be defined as benefits involuntarily received by others and for which they pay nothing (increased benefit payment accuracy). External costs are similarly defined as costs imposed on others without compensation.[5]

Costs and benefits also can be classified as measurable, incommensurable, and intangible. Incommensurable effects are those which cannot readily be translated into the common denominator or denominators that are being used. Intangible effects cannot be measured on any terms. Incommensurable effects are of a material or economic nature, but are simply difficult to measure, such as the recreational value of some jobs. Intangible effects are beyond the economic, such as the preservation of life or freedom. PI studies generally do not have intangible effects.[6]

Developing accurate estimates for the costs and benefits of all program effects is a complex problem, and more often than not results in the development of controversial figures. Even a meticulously prepared CBA is subject to criticism regarding its accuracy. To get some idea regarding how complex a CBA can become, consider what might be involved in developing methods and data for measuring the following: (1) The direct and indirect impacts of all implementation aspects of the corrective action; (2) the lost value of the failure to spend the funding elsewhere; (3) the out-year costs in light of inflationary pressures and expected UI funding; and (4) the incommensurable and intangible effects on worker morale, employers or public opinion.

Table VIII-2 on pages VIII-18 and 19 summarizes some main factors that need to be considered to conduct a CBA. Table VIII-3 on page VIII-20 is a useful worksheet for performing the CBA. When performing a CBA it is important to realize that attention is focused on the economy as a whole, not just the UI agency. A particular action should be recommended only if justified in terms of the CBA.

TABLE VIII-1

95% CONFIDENCE INTERVAL FOR THE EFFECT OF THE BYE TEST PROCEDURE CHANGE IN ERROR RATE

Lower Limit	Point Estimate	Upper Limit
11%	15%	19%

INTERPRETATION

The confidence level is 95 percent that the proposed procedure reduces the error rate for partial payments by between 11 and 19 percentage points. That is, the "chances" are 95 percent that these bounds include the true results for the population.

The confidence level is 97.5 percent that the proposed procedure reduces the error rate for partial payments by at least 11 percentage points. That is, the "chances" are only 2.5 percent that the test procedure reduces the error rate by less than 11 percentage points.

The confidence level is 97.5 percent that the proposed procedure reduces the error rate for partial payments by at most 19 percentage points. That is, the "chances" are only 2.5 percent that the test procedure reduces the error rate by more than 19 percentage points.

Table VIII-2

SOME COST-BENEFIT CONSIDERATIONS

A. Background Issues

1. Are costs and benefits consistently defined?
2. How big a difference will the change make relative to program areas?
3. What are the mathematical formulas for constructing the estimates?
4. Why was this CBA method selected?
5. Is there precedence for using this CBA method?
6. Is there a basis for other SESAs possibly considering the change?

B. Costs

Monetary Costs

1. Have the administrative costs of current error exception handling been estimated?
2. Have trust fund and administrative fund costs been differentiated from operating costs?
3. Has the life cycle of capital been factored into long term costs?
4. Have current and out-year costs associated with varying claims loads and other factors been assessed?

Nonmonetary Costs

1. Will the change adversely affect worker morale by increasing individual workload?
2. Can the change effects be duplicated by average workers rather than the experts used during the study period?
3. Will the change result in loss of duties or prestige for anyone or group?
4. Will the staff responsible for operationalizing the change support the action?

C. Benefits

Monetary Benefits

1. Will a real monetary gain be realized?
2. Is there a necessity for real monetary gain in this case?
3. Do the benefits exist both in the long term and in "worst case" claims load conditions?

Nonmonetary Benefits

1. Is a change justified by its positive effect on claimants, agency staff or employers?
2. Will the change result in agency image enhancement?
3. Will the change prevent future problems?
4. Are there unexpected productivity benefits associated with the change?

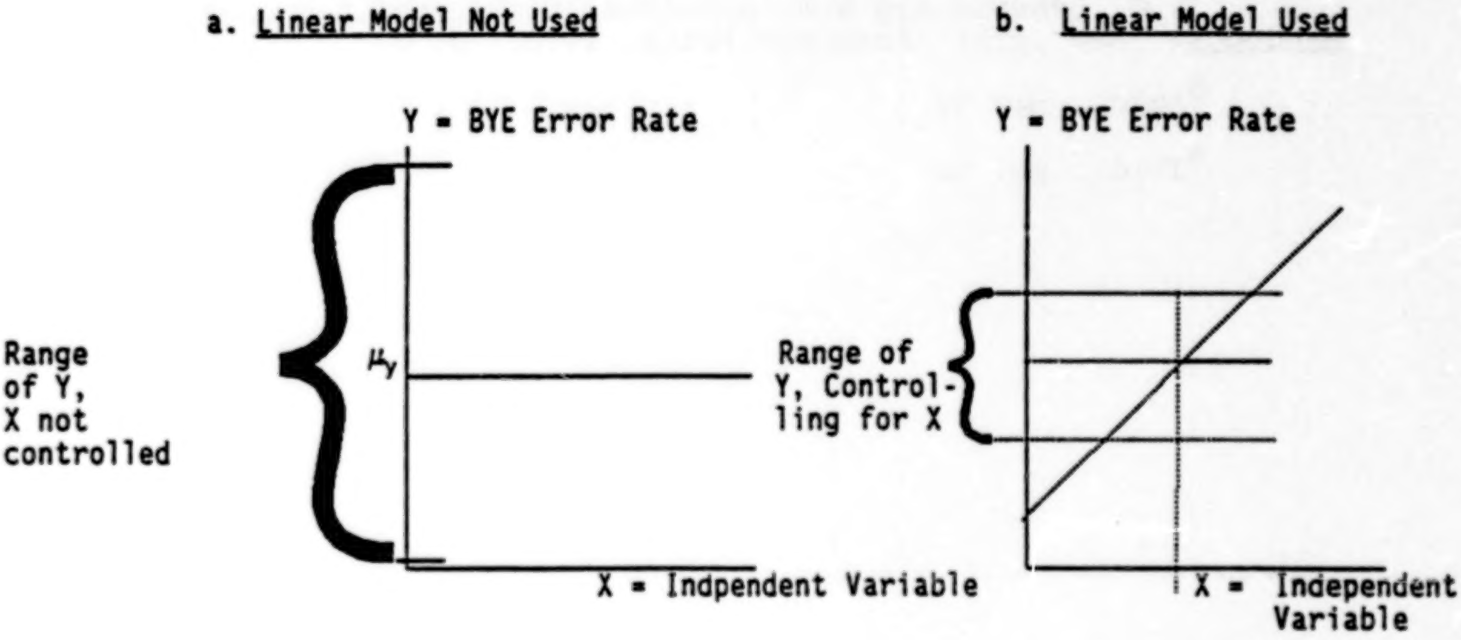
Table VIII-3

CBA Worksheet

		Benefits				Costs			
		Agency	Employer	Claimant	Public	Agency	Employer	Claimant	Public
Monetary Effects	Direct								
	Indirect								
SUM									
Nonmonetary Effects	Direct								
	Indirect								
Intangible Effects	Direct								
	Indirect								

Figure VIII-1

CONFIDENCE INTERVALS FOR BYE EXAMPLE
WITH AND WITHOUT USING THE LINEAR MODEL



Variation Around the Mean of Y

Variation Around the Regression Line for Y

FOOTNOTES

[1] D.K. Hildebrand and L. Ott, Statistical Thinking for Managers. 2nd Edition. Boston: Duxbury Press, 1987. pp. 3-4.

[2] W.S. Cleveland, The Elements of Graphing Data. Monterey, California: Wadsworth Advanced Books and Software, 1985. p. 17.

[3] Ibid. pp. 100-101.

[4] P.G. Sassone and W.A. Schaffer, Cost-Benefit Analysis: A Handbook. New York: Academic Press, 1978. p. 2.

[5] Ibid. pg. 32.

[6] Ibid. pg. 34.

CHAPTER IX

FINAL REPORT

Checklist

1. Have all relevant parties been involved in evaluating study findings and developing recommendations?

   a. ____ The key UI manager(s)?

   b. ____ Other key decision makers?

   c. ____ Relevant staff?

2. Is the final report clear, concise and comprehensive?

   a. ____ Does the report contain a summary of the study problem, purpose and approach as detailed in the proposal?

   b. ____ Are any deviations from the study proposal explained?

   c. ____ Does the report contain a summary of the raw data and a detailed explanation of how it was analyzed?

   d. ____ Are the study findings explained and related to the study objectives?

   e. ____ Were multiple recommendations generated for addressing the problem?

   f. ____ Have cost/benefit considerations been included in a discussion of the recommendations?

   g. ____ Are the operational and administrative implications of the recommendations discussed?

   h. ____ Is some change likely to be implemented as a result of the PI study?

   i. ____ Has a clear and concise executive summary been prepared?

BLANK PAGE

CHAPTER IX

FINAL REPORT

There are no specific format requirements for Program Improvement Study final reports, but the following outline may be useful. Normally the final report should contain most or all of the segments covered below, but it may be organized to best fit individual SESA needs. In any case, reviewing the checklist and suggested format should help ensure that no important areas are omitted.

1. **Cover Page**. The cover page should contain important identifying information. Make the cover page attractive to reflect the quality of the work contained in the report. Many SESAs find it useful to include the following items on the cover page:

- State, agency, bureau and section name;
- Names of key staff and project manager;
- Title of PI study;
- Period covered by study;
- Date report completed, month and year; and
- Graphics - State seal and agency logo.

2. **Executive Summary**. The executive summary is an important part of any final report. It is designed for key staff who do not have time to read the entire report. Key issues to address include background and study purpose, major findings, conclusions, recommendations and any PI action taken or anticipated. Although the executive summary is the first section of the final report, it normally should be the last item written. It should be brief, to the point and limited to a few pages in

length. A suggested outline for the executive summary is included in Table IX-1.

3. Background/Problem Assessment. This section of the report should fully describe how and why the project was initiated. Include important details concerning goals, objectives and what staff were involved in the study. Much of the content for this section can be gleaned from original agency correspondence and the study project proposal. Further, preparation of this section provides an excellent opportunity to document administrative or technical difficulties that occurred during the study period. Addressing these occurrences early in the report may be better than detailing them in the midst of the later technical discussions. A suggested outline for background/problem assessment issues is provided in Table IX-1.

4. Description of Study Approach and Data Collection. Considerable detail in this section is desirable. The methodology should be explained as it actually was employed, and the actual study approach should be compared to the approach contained in the original proposal if the two differ. Unforeseen circumstances and constraints often result in methodological changes during the study period. Such changes are common, but they should be explained. Make the technical sections of the text as understandable as possible. Samples of survey instruments and other such materials normally would be relegated to the Appendices. A possible outline for study approach and data collection is provided in Table IX-1.

5. Analysis and Findings. This section of the final report deals directly with study results and findings. In assessing study results, it is important to document any deviation from the

proposed analysis plan, and to clearly establish how the quantitative or qualitative analysis actually was conducted.

Graphs, charts and tables are very useful in reinforcing important relationships and key study findings. But keep them simple and well labeled. If survey questionnaires were utilized, they should be included as an appendix. Keep the narrative focused on the relationships between study objectives and findings. Involved tables and volumes of raw data should be placed in appendixes rather than the body of the report. A suggested outline for analysis and findings is provided in Table IX-1 on pages IX-6 and 7.

6. **Conclusions and Recommendations**. This section of the final report is essential in conveying the knowledge gained from the project. A logical scheme for presenting recommendations is to use a rank ordering based upon the cost/benefit analysis. Clearly explain the basis for the rankings to clarify what factors resulted in the particular ranking reported. It is essential to explain the validity and reliability of the findings, as well as the degree they can be generalized to the population. A clear distinction should be made between conclusions that were the primary focus of the study and those unexpected or casual findings that usually turn up during the course of a study. This is done to prevent confusion about recommendations and conclusions seemingly unrelated to the purpose of the study. The conclusions and recommendations should be specific. They should clearly state what can be done, what impact the action will have and why. Avoid recommendations not supported by the study. For example, the key recommendation of a PI study might be simply to test two alternatives in an operational environment. If this is the case, it is necessary to

fully develop and explain the alternative procedure(s) that would be tested, along with expected impacts on the problem.

The final key to an effective final report has been addressed throughout the TAG, the need to keep key staff and primary parties fully involved in all phases of the study. That is, developing and maintaining organizational support is crucial from start to finish. This can be critical when analyzing the data and developing conclusions and recommendations. Administrators do not appreciate "bombshell" recommendations that suddenly appear in a final report. Likewise, program area staff should be provided the opportunity for full participation in developing study conclusions and recommendations. Their intimate knowledge of the study area will provide valuable insights for data interpretation as well as for preparing realistic and workable recommendations. By actively seeking their input, their support for any changes that may be recommended is much more likely. In contrast, developing recommendations in a vacuum often results in a complete breakdown of the study at its most critical point. Even the best recommendations can be scuttled without organizational commitment and support. A suggested outline for conclusions and recommendations is included in Table IX-1.

TABLE IX-1

ILLUSTRATIVE FINAL STUDY REPORT CHECKLIST

A. Executive Summary

1. Why was the study undertaken?
2. What specific area of the UI system was studied?
3. What are the primary findings?
4. What specific program changes do the findings appear to support?
5. What are the major limitations of the study findings?
6. What are the major conclusions and recommendations?

B. Background and Problem Assessment

1. Why was the study initiated?
2. What problem assessment was conducted?
3. What level of administrative support was achieved?
4. What groups were consulted during the problem identification process?
5. What data were used to isolate the problem?

C. Study Goals and Objectives

1. What was the primary goal of the study?
2. What were the primary objectives for meeting the goal(s) of the study?
3. State any secondary goals and objectives, if any.
4. How were the study goals and objectives realized?
5. Was the purpose of the study to collect additional data or test an alternative procedure?

D. Study Design and Data Collection

1. What basic design was selected?
2. Why was this approach selected?
3. Are the limitations of the design discussed?
4. How was implementation of the study design monitored and controlled?
5. What controls were constructed to address foreseen or unforeseen complications?
6. Were there deviations from the design?
7. Describe and explain the effects of such deviations.
8. How were the data collected?
9. Why were the data collected in this manner?
10. Who conducted the data collection?
11. Were the data collection techniques and tools pretested?

E. Analysis of Findings

1. Were any participants not included in the analysis? If so, why?
2. What statistical techniques were used in the data analysis?
3. Were the results determined to be statistically significant?
4. Can the results be extrapolated to the UI population?
5. Are the results practically important for policy purposes?
6. Do any graphs or charts accurately reflect the study findings?
7. Are the limitations of the findings clearly stated?
8. If appropriate, have confidence intervals been included to place bounds on estimate?

F. Conclusions and Recommendations

1. Are the conclusions and recommendations clearly but succinctly stated?
2. Is the basis (cost/benefit analysis) for ranking recommended changes clearly explained?
3. Are the organizational support issues involved in conclusions and recommendations addressed?

CHAPTER X

IMPLEMENTATION AND FOLLOW-UP

Checklist

1. Has a comprehensive program improvement implementation plan been developed?

    a. ____ Have key staff, program managers and operational staff been involved in development of the workplan?

    b. ____ Have all policies, procedures and processing activities that require change been identified?

    c. ____ Have realistic time frames for training operational staff been established?

    d. ____ Are sufficient technical staff available to provide assistance to bring about the change(s)?

    e. ____ Has the impact on the Data Processing System and other affected operations been evaluated?

    f. ____ Has a method been developed for keeping affected parties informed of progress towards implementation?

2. Have significant implementation barriers been identified?

    a. ____ Have any technical barriers that would hinder or prevent implementation been addressed?

    b. ____ Have any operational barriers which would hinder or prevent implementation been addressed?

    c. ____ Have appropriate staff with insight on identifying or resolving these barriers been consulted?

3. Has a plan for evaluating the impact of the action been developed?

    a. ____ Does the plan require modification of existing evaluation techniques to collect required information?

b. ____ Has an adequate learning period been incorporated into the evaluation plan?

c. ____ Has a particular person or unit been assigned responsibility for preparing the evaluation findings report?

CHAPTER X

IMPLEMENTATION AND FOLLOW-UP

The final stages of any PI study involve implementation of feasible program improvement actions and an assessment of the impact of these actions on the UI system. Achievement of these objectives requires development of a plan that involves gaining the support from key agency staff, development of a comprehensive implementation plan and development of a strategy for assessing the practical effect of the change.

1. __Implementation Plan__. SESA administrators and staff have substantial experience in implementing complex UI law changes and developing innovative automated systems. For this reason, no attempt is made to elaborate on the implementation plan items found in item 1. of the checklist. Nonetheless, it may be helpful at this stage of the implementation process to consult the checklist to be sure critical steps in developing the implementation plan have not been overlooked. Additionally, if workplan charts have not been a part of past implementation efforts, it may be useful to consider preparing workplan charts similar to those found in Chapter V.

2. __Potential Barriers__. It is important to identify the potential barriers that may be encountered during implementation of PI study recommendations. In addressing the relative importance of various barriers, it may be helpful to categorize them by type. For purposes of this discussion, potential barriers are classified as either technical or operational.

__Technical Barriers__ - Technical barriers are defined in this context to include such restraints as legal impediments, political considerations, regulatory impediments, data processing system limitations, or the lack of available funding. Technical barriers are generally broad in scope, and any of them could prevent the agency from implementing a PI change.

Operational Barriers - Operational barriers are less involved restraints that are usually identified when assessing the procedural modifications necessary to implement a change. These barriers may include the effects of the change on payment timeliness, associated payment or charging functions, the maintenance of required reporting information, or even employee morale.

3. **Identifying Barriers**. It usually is a rather simple process to identify the major technical barriers to implementing a particular program change. Such constraints are generally obvious to experienced SESA staff and should be considered and resolved as the first step in planning an implementation. In contrast, operational barriers revealed during procedural development often are more subtle and less obvious, but they must be emphasized in considering the practical application of the change. The completion of the overall evaluation of potential barriers requires significant coordination with those responsible for data processing systems, UI program technical support areas and operational areas. Involving a wide variety of staff will ensure that consideration is given to any unanticipated side effects of the action. This is a necessary process because the UI payment process is an integrated system rather than a series of isolated activities. Even an apparently minor alteration of the process can result in large ripple effects.

Positive ripple or side effects, even if unanticipated, present no problem for PI study changes, so the main interest in assessing side effects usually is on potentially negative side effects. Examples of such effects include claim certification processing delays that degrade timely payment performance, loss of information critical to completing SESA or federally required reports, or the need to alter methods for handling exception processing at the various decision points in the UI claims process. SESAs generally have established practices for evaluating potential program modifications, such as construction of detailed flow charts of the new procedure, round table discussions with affected staff, preparation of a detailed workplan, and operational testing of the program change at selected worksites. Utilizing these well known techniques will help keep the implementation process on track.

4. Assessing the Impact of the Action. Once a program improvement action has been taken, it is necessary to develop a strategy for assessing its positive or negative impacts on the system. Because it is unlikely that SESAs will undertake a full-blown follow-up study of the program change, it is recommended that consideration be given to identifying methods for tracking results at the time a decision is made to implement the action. In this way, it is possible to build a reporting system that will track the effects of the change from the very beginning. This will be helpful in evaluating the effectiveness of the implementation plan (training, forms etc.) and will expedite data collection and analysis when assessing the impacts of the change. There are many system evaluation programs within every SESA, including BQC, QA and LO reviews. It is extremely likely that these systems already are used to evaluate the program areas affected by PI study changes. It may be possible

143

to collect additional data during these evaluations to assist in evaluating the effectiveness of the PI change. Augmenting existing evaluation systems will reduce the costs associated with evaluating the program area and ensure that the impact of the action has been adequately assessed.

The final step in establishing an adequate method for assessing the program change is to specifically assign responsibility for preparing impact reports to a single individual or unit. This is useful because the data required for this assessment likely must be gathered from multiple sources. As such, the process must be coordinated in a manner similar to the problem identification phase of study development.

APPENDIX A

CONTENTS

Page

Unemployment Insurance Program Letter No. 37-91:
Guidelines for Unemployment Insurance (UI)
Quality Control (QC) Program Improvement (PI)
Studies . . (August 13, 1991). A-1

Bibliography. A-8

Glossary. A-10

BLANK PAGE

U.S. DEPARTMENT OF LABOR
Employment and Training Administration
Washington, D.C. 20210

CLASSIFICATION
UI/QC/PI Studies

CORRESPONDENCE SYMBOL
TEUQI

DATE
August 13, 1991

DIRECTIVE : UNEMPLOYMENT INSURANCE PROGRAM LETTER NO. 37-91

TO : ALL STATE EMPLOYMENT SECURITY AGENCIES

FROM : DONALD J. KULICK
Administrator
for Regional Management

SUBJECT : Guidelines for Unemployment Insurance (UI) Quality Control (QC) Program Improvement (PI) Studies

1. Purpose. To provide revised and expanded guidance to State employment security agencies (SESAs) on the development and conduct of special studies that may facilitate UI program improvement actions.

2. References. UIPL 29-86 (April 18, 1986); and FM 11-90, (October 27, 1989).

3. Background. The Employment and Training Administration (ETA) has always encouraged the States to develop innovative and cost effective methods for obtaining and using benefits QC (BQC) data. The QC regulation at 20 CFR 602.21(b) -- following the consensus agreement -- provides a degree of flexibility in the use of BQC resources for basic study activity. In December 1987 the Department issued a directive (UIPL 9-88) to the SESAs confirming their option to divert QC staff resources, for an approved period, to conduct program improvement (PI) studies, with a corresponding reduction in the QC weekly samples during the study period. The directive stressed that such reductions in QC sampling were exceptional and temporary, and that diversion of QC investigative staff for a variety of special studies should be approved in advance.

UIPL 33-89 offered broadened guidelines to encourage SESAs to study a wide range of UI program issues and problems. Since its issuance, a large number of States have conducted PI studies. The quality level of these studies varied significantly. That experience is the basis of a number of modifications and refinements of existing PI study policy and procedures incorporated in this program letter. These

RESCISSIONS	EXPIRATION DATE
UIPL 33-89	August 31, 1992

DISTRIBUTION

changes are intended to enhance the general quality of PI study proposals submitted for approval, and to improve the processing of approvals of both SESA proposals and requests for study extensions by the Regional and National Offices (RO and NO).

4. Policy. This directive reiterates the ETA policy of encouraging and aiding SESAs to undertake short-term research projects and program studies as a means of translating general findings of QC case investigations into practical UI program improvements and error reductions. Such efforts may involve UI data analysis beyond the capabilities of the QC analyst and other available UI staff, or take the form of gathering new, supplemental data that could serve to inform management policy and process decisions.

No approval is required when the State maintains its required QC sample size and funds the project from available resources. Advance approval from the NO is required if a State plans to fund a study by reducing its required annual QC sample, and deploys QC staff in this study effort. No State may reduce its QC sample size below 400 cases per year.

a. Required Standards. Standards for SESA PI study proposals should reflect primarily the needs of SESA decisionmakers. Areas of study may include development and testing, but not implementation or operation of ongoing UI activities. Therefore, QC staff cannot be diverted from QC functions to work on regular UI program implementation or operational activity.

During the past three years, SESAs have submitted program improvement proposals which varied widely in content and quality. This variance may have been the result of broadly written guidelines. It is now appropriate to more specifically define the basic content requirements for study proposals.

(1) Problem Addressed - What is the specific problem or UI program facet that is proposed for PI study? What aspects of this situation have SESA staff identified through QC verification data? Indicate the findings that SESA staff have made through analysis of QC data (and other UI data) that suggest the need for further study before SESA management can consider potential program improvement actions.

(2) Objectives and Methodology - What are the study objectives? What does the SESA expect to learn from the study? Describe the methodology to be followed. Explain how this study approach will provide additional information needed by UI management.

(3) Results Expected - What results are anticipated? How will the results of the study be presented to UI principals? How does the SESA anticipate translating this information into decisions that may lead to UI program improvement?

(4) Projected Study Resources - How many QC staff and other UI resources will be needed to conduct the program improvement study? Be sure to identify QC positions the SESA plans to dedicate to the study and indicate the size (number of cases) of the reduction in QC weekly sampling requested during the study period.

(5) Work Plan - What is the projected time period for: (a) implementation of the PI study, and (b) preparation of a final study report? Present a study work plan that specifies: (a) QC staff hours and time (in weeks) required for completing major tasks or segments of the study, and (b) QC staff time, if any, needed for preparation of the final report. Indicate the: (a) anticipated study start date, (b) study completion date, and (c) planned date for completing the final report.

In short, by responding to the several questions and issues presented above, the SESA should be able to communicate clearly the following essential information in each study proposal offered:

- the problem or issue to be studied,
- the objectives of the study,
- the methodology to be followed,
- the results and outcomes expected and how to use these findings for program refinement,
- UI resources (including QC staff) to be used in completing the study and the final report, and the schedule for completion of the study and the final report.

In the preparation of PI study proposals, SESA staff may find it useful to organize its material following an outline which details the basic study elements and information that are required in this directive. An example of such a proposal outline or checklist is presented in the attachment.

If a study proposal fails to meet the five established requirements, the RO will return it to the SESA for completion, and for subsequent resubmission to the RO. No PI study involving QC sample reduction should be launched prior to NO approval. The NO will not approve such studies retroactively.

b. Final Report - At the conclusion of the QC program improvement study, the SESA should prepare a final report assessing the outcomes of the study. An adequate study report will cover the following elements:

(1) the focus and scope of the study, i.e., the problem, goals and objectives of the proposed study;

(2) the methodology pursued and an assessment of whether or not the study objectives were met by this approach.

(3) a brief description of study implementation and the outcomes (expected, unexpected) of the effort;

(4) a statement of the principal findings (positive or negative) and the potential usefulness of these results for prompting UI program improvements; and

(5) recommendations for UI management consideration. Describe any benefits that may have resulted already from the study.

The SESA should forward four copies of each study report to the appropriate RO. (It is desirable that, whenever feasible, the SESA send final reports of completed studies to the Region along with new and subsequent study proposals that are submitted for approval.)

5. Approval Procedures for Study Proposals. SESAs wishing to employ QC resources in UI program improvement studies will prepare proposals for submittal to their respective Regions, following the guidance and requirements of this directive. Since approvals of QC staff diversions are directly related to the Fiscal Year (FY), it is highly desirable that State proposals be planned and submitted to the Regions on a FY basis. (Several SESAs, particularly those using the Quality

Improvement Project study model, have developed PI study proposals that coincide with the FY budget cycle.)

SESAs are encouraged to seek assistance from the Regions, when needed, in developing suitable study work plans for inclusion in a request for QC staff diversion. The following steps are intended to expedite the review and approval of PI study proposals by the RO and NO.

a. The SESA should submit a complete study proposal to the RO two months prior to the planned start date of the proposed study. This timing is particularly important in situations in which a SESA is hopeful of securing approval for additional studies to be launched following one or more current studies then nearing completion.

b. The SESA should indicate the completion date of the final report and specify QC resources needed, if any, and the schedule of sample reduction requested for completion of the report.

c. Regional staff will review the proposal and, if it meets the requirements stipulated in Section 4. a. above, send it along with the RO's evaluation and recommendations to the NO for review and approval. This review will be completed by the Region in three weeks or less.

The NO decisions will be provided to the Region within the remaining five weeks.

6. Requests for Study Extensions. The NO policy continues to allow the diversion of SESA QC staff, with accompanying sample reductions, to accomplish UI program improvement as long as the State is engaged in activities likely to produce positive results. The policy is as follows:

a. States that complete their PI studies in a timely manner may submit a new proposal at any time, in accordance with policy and procedures spelled out in this directive. In the interim they must return to investigating their respective full weekly sampling allocations.

b. A State that has not completed one or more studies in the approved study period may request an extension of the period of QC staff diversion and sample reduction.

151

The written request should be submitted to the appropriate RO as soon as it becomes clear that the study in question will not be completed as originally approved. At a minimum, such extension requests should be forwarded to the RO at least 45 days prior to the expiration of the initial approved studyperiod. For projects approved on Fiscal Year basis, this means submission of requests on or before August 15.

The SESA request for extension should include the following information: (a) explanation for the extension request; (b) a supplemental work plan which details the number of QC staff and the number of weeks required for completion of both the study and the final study report, and (c) proposed RO assistance, if needed.

The RO will forward the request to the NO, with its recommendation, 30 days prior to the end date of the ongoing study period, so that the NO is able to review the request and send a response to the Region before expiration of the initial approval period. If there is a delay beyond the initial study expiration date in approval of such an extension request, the State must return to its regular weekly sample level until approval has been confirmed.

7. Action Required. SESA administrators are requested to provide this information to appropriate SESA staff.

8. Inquiries. Questions should be directed to the appropriate RO.

9. Attachment. Minimum Proposal Elements/Requirements.

Minimum Proposal Elements/Requirements
(Address each item)

A. Problem Identification

- (1) Statement of problem
- (2) BQC data pointing to the problem
- (3) Other pertinent UI data analyses

B. Objectives and Methodology

- (1) Study objectives
- (2) Expected insights from the study
- (3) Methodology/study approach
- (4) Potential value of study for UI management

C. Potential Results

- (1) Results expected
- (2) Presentation of findings to management
- (3) Likely program improvement actions

D. Study Resources

- (1) Staff requirements
- (2) QC sample reduction needed

E. Work Plan

- (1) Schedule of key activities
- (2) Deployment of QC staff
- (3) Study completion date; final report date

BLANK PAGE

BIBLIOGRAPHY

Backstrom, C.H. and Hurch-Cesar, G. Survey Research. New York: John Wiley and Sons, 1981.

Bedworth, D.D. Industrial Systems: Planning, Analysis, Control. New York: The Ronald Press Company, 1973.

Blalock, A.B. Evaluating Social Programs at the State and Local Level. Kalamazoo, Michigan: Upjohn Institute, 1990.

Blalock, H. Social Statistics. New York: McGraw-Hill, 1979.

Bowen, B.D. and Weisberg, H.F. An Introduction to Data Analysis. San Francisco: W.H. Freeman, 1977.

Campbell, D.T. and Stanley, J.C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally College Publishing Company, 1963.

Champers, J.M., W.S. Cleveland, B. Kleiner and P.A. Tukey. Graphical Methods for Data Analysis. Boston: Duxbury Press, 1983.

Cleland, U.L. Project Management: Strategic Design and Implementation. Blue Ridge Summit, PA: TAB Books, 1989.

Cleveland, W.S. The Elements of Graphing Data. Monterey, California: Wadsworth Advanced Books and Software, 1985.

Cochran, W.G. Planning and Analysis of Observational Studies. New York: John Wiley & Sons, 1983.

Dekom, A.K. Practical Project Management. 1st ed. New York: Random House Business Division, c. 1986.

Dillman, D.A. Mail and Telephone Surveys: The Total Design Method. New York: John Wiley & Sons, 1978.

DuBrin, A.J., R.D. Ireland and C.J. Williams. Management and Organization. 6th ed. Cincinnati: South-Western, 1989.

Duncan, A.J. Quality Control and Industrial Statistics. 5th Edition. Homewood, Illinois: Richard D. Irwin, 1986.

Frame, J.D. Managing Projects in Organizations. 1st ed. San Francisco: Jossey-Bass, 1987.

Graham, R.J. Project Management: Combining Technical and Behavioral Approaches for Effective Implementation. New York: Van Nostrand Reinhold, c. 1985.

Hildebrand, D.K. and L. Ott. Statistical Thinking for Managers. 2nd Edition. Boston: Duxbury Press, 1987.

Kerlinger, F.N. Foundations of Behavioral Research. 3rd Edition. New York: Holt, Rinehart and Winston, 1986.

Kirk, R.E. Experimental Design: Procedures for the Behavioral Sciences. 2nd Edition. Belmont, California: Wadworth, 1982.

Kish, L. Survey Sampling. New York: John Wiley & Sons, 1965.

Kish, L. Statistical Design for Research. New York: John Wiley & Sons, 1987.

Kossen, S. The Human Side of Organizations. 5th Edition. New York: Harper-Collins, 1991.

Lansing, J.B. and J.N. Morgan. Economic Survey Methods. Ann Arbor, Michigan: Institute for Social Research, Survey Research Center, 1971.

Mishan, E.J. Cost-Benefit Analysis. New York: Praeger Publishers, Praeger Special Studies in International Business, Finance and Trade, 1976.

Neter, J., W. Wasserman and M.H. Kutner. Applied Linear Regression Models. 2nd Edition. Homewood, Illinois: Richard D. Irwin, 1989.

O'Sullivan, E. and G.R. Rassel. Research Methods for Public Administrators. New York: Longman, 1989.

Poister, T.H. Public Program Analysis: Applied Research Methods. University Park: University Park Press, 1978.

Richmond, S.B. Operations Research for Management Decisions. New York: The Ronald Press Company, 1968.

Rossi, P.H. and H. Freeman. Evaluation: A Systematic Approach. Newbury Park: Sage, 1985.

Sassone, P.G. and W.A. Schaffer. Cost-Benefit Analysis: A Handbook. New York: Academic Press, 1978.

Scheaffer, R.L., W. Mendenhall, and L. Ott. Elementary Survey Sampling. 3rd Edition. Boston: Duxbury Press, 1986.

Shannon, R.E. Systems Simulation: The Art and Science. New Jersey: Prentice-Hall, 1975.

Sudman, S. and N.B. Bradburn. Asking Questions: A Practical Guide to Questionnaire Design. San Francisco: Jossey-Bass, 1989.

156

GLOSSARY

AGGREGATE DATA - Data relating to specific groups such as cities, counties or claimants that are the result of a measure on smaller units within the larger group.

ANALYSIS OF VARIANCE - A statistical test for analyzing data from experiments. Especially useful if the experiment has more than one independent variable or more than two kinds of an independent variable.

ARTIFICIAL EFFECTS - See Hawthorne effect.

BIAS - A systematic difference between a sample statistic and the population parameter it is to measure because of a flaw in the sampling design, its implementation, or the data-collection procedure. The sample either systematically under- or overestimates the population parameter.

BIASED QUESTION - A survey question that elicits inaccurate information because it is worded in a manner that encourages respondents to give one answer rather than another.

BQC SYSTEM - Benefits Quality Control program which systematically investigates samples of the population of UI payments each week.

CASE STUDY - A type of study in which a person, program, agency or some other unit of analysis is examined in detail. Comparisons are not usually made with other units having different values of the variables of interest.

CAUSE VARIABLES - See independent variable.

CENSUS - An accurate count of an entire population.

CETERIS PARIBUS (L.) - Holding all variables, except the ones you are interested in constant: "everything else being equal."

CLUSTER SAMPLE - Probability sampling in which groups or jurisdictions comprising groups of units are randomly selected for a sample.

CONFIDENCE INTERVAL - The interval placed around the value of a sample statistic in which the investigator expects the value of the corresponding population parameter to be located.

CONFIDENCE LEVEL - The confidence that an investigator can have that a sample estimate is within a specified range of the population parameter.

CONFOUNDED EFFECTS - Observed conclusions which may not apply to other treatment groups because the conclusions are the result of the specific mix of planned and unplanned stimuli.

CONFOUNDING VARIABLES - Variables that are unintentionally manipulated. If confounding variables are present in a study, it is hard to say what caused an observed effect.

CONTROL GROUP - Subjects who do not receive the experimental treatment. These subjects are compared to the treatment group to determine whether the treatment had an effect.

CONTROL VARIABLE - A variable included in an analysis to determine whether it affects the relationship between two other variables.

CONTROLLED VARIABLES - Extraneous variables that can be controlled adequately by the research design.

DATA MATRIX - A table developed to assist the PI study planner in evaluating the relevance of various data sources to the area of potential study.

DECISION ANALYSIS	Statistical analysis based on Bayesian approach to statistics. Introduces the payoff concept to statistical analysis.
DEPENDENT VARIABLE	The factor that the experimenter predicts is affected by the independent variable. The response that the experimenter is measuring.
DISTURBING VARIABLES	An extraneous variable which may be confounded with dependent and independent variables. Failure to either control or randomize these variables is the primary problem with non-experimental designs.
ENVIRONMENTAL VARIABLES	See Extraneous Variables.
EXPERIMENTAL CONTROL	The process of controlling the effects of extraneous variables by using a rigorous experimental design.
EXPERIMENTAL DESIGN	A design in which subjects are randomly assigned to be exposed to a treatment and compared to other subjects, also randomly assigned, who are not exposed to the same treatment.
EXPLANATORY VARIABLES	Also known as experimental variables. These variables are the objects of the research design. Comprised of two variable types: independent and dependent variables.
EXPLORATORY STUDIES	A type of field study designed to identify possible relationships between variables and to identify possible hypothesis to provide the framework for more rigorous study.
EXTERNAL VALIDITY	The appropriateness of extending or generalizing findings to a group beyond those in the study.

EXTERNAL VALIDITY THREATS - Factors that limit the ability to apply the findings of a study to cases not involved in the study. These factors include: unique program features, effects of selection, effects of setting, effects of history, effects of testing, and reactive effects of experimental arrangements.

EXTRANEOUS VARIABLES - Variables which may interfere with the relationship between dependent and independent variables, thereby biasing study findings.

FIELD EXPERIMENT - An experiment performed in a nonlaboratory setting.

FIELD STUDY - Non-experimental studies conducted in a natural setting and aimed at discovering the relationships between variables of interest. Case studies are the most common example of field studies.

GANTT CHART - A method for developing work plans which involves charting a study's major activities and the expected dates for completing the activities.

GENERALIZABILITY - The extent to which a study's findings can be extrapolated to the sample and target populations.

GOAL - A general statement of a desired end product.

HAWTHORNE EFFECT - The effect of any intervention, including the study itself upon the actions of those being observed. This effect can occur in any study regardless of specific design.

HYPOTHESIS - A testable prediction.

INFERENTIAL STATISTICS - The development and use of probabilities to draw conclusions about a given population on the basis of evidence from a sample.

INTERNAL VALIDITY	- The extent to which a design provides evidence that a specific independent variable caused a change in a dependent variable.
INTERNAL VALIDITY THREATS	- A series of factors that could be the cause of a change in a dependent variable and represent alternatives to the independent variable of interest. These threats include history, maturation, statistical regression, selection, experimental mortality, testing, instrumentation and design contamination.
LABORATORY EXPERIMENT	- An experiment in which it is possible to control the effects of nearly all extraneous variables and to directly test the relationship between independent and dependent variables.
MATCHING	- Construction of control groups by finding subjects identical in relevant respects to subjects in the test group.
NEEDS ASSESSMENT	- A systematic appraisal of type, depth and scope of a problem.
NON-EXPERIMENTAL DESIGNS	- Designs which consist solely of the observation of naturally occurring events without any attempt to manipulate their occurrence.
NONPROBABILITY SAMPLE	- Samples of a population that are made based upon judgment or other factors of interest not related to known probabilities.
NONRESPONSE BIAS	- A type of nonsampling error which results from the failure to measure some of the sampled units.
NONSAMPLING ERRORS	- Threats to the validity of a study which result from non-response, measurement error and coding error.

OBJECTIVE	-	Specific, operationalized statements detailing the desired accomplishment of a study.
OUTCOME VARIABLE	-	See Dependent Variable.
OUT-YEAR COSTS	-	The increased costs of an action over time due to many factors including increasing wages, inflation or equipment life cycle.
PARAMETER	-	A characteristic of a population. Analysts draw samples to estimate the values of population parameters.
PERT CHARTS	-	All extensions to GANTT charts that are used to determine the probability that the study/project will be completed within a particular time frame.
PILOT TEST	-	A small study designed to test the adequacy of a proposed data collection strategy.
POPULATION	-	A total set of units or items sharing at least one characteristic.
POPULATION IN NEED	-	Units of a given population that currently manifest a given condition.
POPULATION AT RISK	-	Units of a given population that have a significant probability of developing a given condition.
POSTTEST ONLY CONTROL GROUP	-	An experimental design in which test and control groups are randomly selected for exposure/non-exposure to a treatment, and post-treatment scores (but not pre-treatment) are obtained to evaluate the effect of the treatment.
PRACTICAL SIGNIFICANCE	-	The extent to which a statistical finding is important.
PRECISION LEVEL	-	The expected range of error produced by a measure.

PRIMARY PARTIES - Individuals or groups with a direct stake in the study process or outcome.

PROBABILITY SAMPLE - A type of sampling in which each unit of the population has a known, nonzero chance of being selected.

PROBLEM IDENTIFICATION - The comprehensive evaluation and analysis of BQC and other data within a SESA; the first step in evaluating the need for a PI study.

QUASI-EXPERIMENTAL DESIGN - A design having some but not all of the characteristics of the experimental design; used in applied research studies when the control necessary for a true experiment is not possible or practical.

RANDOMIZED VARIABLES - Uncontrolled extraneous variables, the net effect of which are limited by randomly assigning subjects to test and control groups.

REGRESSION TO THE MEAN - Also statistical regression. Refers to the likelihood that on any given observation some cases take on extreme values which deviate considerably from their normal range. These cases will tend to "regress" to their normal values on subsequent observations.

RELIABILITY - The extent to which obtained statistics are reproducible in repeated administration of the same treatment. Refers to the consistency of a measurement tool and the degree to which scores are free from random error.

REPLICATION - The repeated testing of a specific treatment to establish the generalizability of the treatment.

SAMPLE - A subset of a population that is analyzed for the purpose of gaining insight about the entire population.

SAMPLED POPULATION - The population from which a sample is selected.

SECONDARY PARTIES - Individuals or groups not directly affected by the study process or outcome.

SELECTION BIAS - Apparent treatment effects that are due to comparing groups that differed before the treatment was administered (comparing apples to oranges).

STATISTIC - A characteristic or value of a characteristic of a sample. Investigators use sample statistics to estimate population parameters.

STATISTICAL CONTROL - Statistical techniques that control differences between test and control groups.

STATISTICAL SIGNIFICANCE - Confidence beyond a reasonable doubt (usually 95%) that observed relationship was not due to random error, but represents a real relationship.

SURVEY RESEARCH - A type of exploratory field study that focuses on the accurate assessment of the characteristics of whole populations of people on the basis of a sample.

SYSTEMATIC SAMPLE - Probability sample that is based upon a skip interval from a randomly determined starting point.

TARGET POPULATION - The set of units to which the study planner desires to apply the study results.

TEST GROUP - A group of subjects administered a treatment which is later compared to a control group of untreated subjects.

TIME SERIES DESIGN - A research design in which the effects are estimated from a series of repeated measurements taken before and after an intervention.

TREATMENT GROUP - See Test Group.

UNIVARIATE ANALYSIS - Descriptive analysis of the distribution of one variable.

VARIANCE - A measure of dispersion calculated by squaring the difference between the mean and the value of each case, summing these values and dividing by the total number of cases. The variance is the square of the standard deviation.

BLANK PAGE

APPENDIX B

CONTENTS

Page

Sample Cover Letter to Regional Office Regarding PI Study Proposal. B-1

Sample Program Improvement Study Proposal: Benefit Year Earnings. B-2

Facsimile Letter to Regional Office Regarding PI Study Proposal: Benefit Rights Interview (Outline) B-16

Abbreviated Sample Program Improvement Study Proposal Benefit Rights Interview B-17

BLANK PAGE

NAME
Regional Administrator
U.S. Department of Labor
Employment and Training Administration
STREET
CITY, STATE ZIP

Dear :

We have been working in conjunction with members of your staff to develop a Quality Control Program Improvement Study. I am pleased to inform you that we have completed a proposal for a study of our Benefit Year Earnings procedures. This proposal (enclosed) is being submitted for your review and submission to the National Office.

I feel this study will provide valuable information from which we will be able to make an informed decision regarding the best method of verifying such earnings. If the study reveals that a procedural change is warranted, it would be a major step forward in improving our UI system and service to our unemployed workers.

Please feel free to contact me at your convenience, if additional information regarding this study is necessary.

Sincerely,

SESA Administrator

Enclosure

BLANK PAGE

(State) Department of Labor & Industry
Bureau of Unemployment Compensation
Quality Control Section

SAMPLE
PROGRAM IMPROVEMENT STUDY
PROPOSAL

BENEFIT YEAR EARNINGS:

Test and Control Group Study
of an
Alternative Earnings Reporting Method

Reference - UIPL 37-91 (August 13, 1991)
- Technical Assistance Guide

171

BLANK PAGE

A. PROBLEM IDENTIFICATION

1. **Statement of Problem/Background** - Our state currently verifies all benefit year earnings using a mail verification procedure. Claimants report partial earnings when signing for a week of benefits, and the agency mails a form to the employer. The employer has five (5) working days to return the form with a breakdown of earnings and missed hours for the claim week. If the earnings report is not returned within the allotted time, the payment is made based upon the claimant's estimate of earnings. Because employers frequently fail to return the form in the required time period, benefits are frequently paid based solely upon the claimant's estimate of earnings. Claimants frequently report earnings incorrectly, resulting in mispaid benefits. Although many of these mispayments are corrected when the form is returned (late), the process of correcting the mispayment requires expenditure of considerable operational resources.

Field office managers and regional staff were consulted during regular managers' meetings for input regarding the problem. In addition, state office program staff held several meetings to review field input and additional information obtained from other state agencies with respect to benefit year earning procedures.

Our corrective action committee has reviewed the available data on this issue. The committee concluded that additional information was necessary before taking any action to change the current procedures. The following alternatives for a PI study were discussed at length.

a. Conduct a PI study to test the effect of providing an information packet to all employers informing them of their responsibility to return earnings verifications timely.

b. Conduct a PI study to test the effectiveness of a new earnings verification form.

c. Conduct a PI study to test an alternative method of securing timely and accurate earnings verification from employers.

After careful consideration of the information available and the input from field and program staff, and with the concurrence of the appropriate SESA administrators, we have

BLANK PAGE

concluded that the third option (c) will best serve our requirements for establishing both the extent of the problem and providing insight regarding a possible solution.

We believe that the benefit year earnings verification process can be improved by requiring claimants to bring a completed earnings verification form, signed by the employer, to the local office when filing a partial claim. By modifying the BYE reporting requirements in this manner we hope to achieve the dual benefit of: (1) reducing BYE mispayments, and (2) reducing operational costs associated with correcting payments when employers return verification forms after the allotted time period. Prior to implementing the procedure statewide, however, we would like to conduct a PI study to ascertain the effectiveness of the two procedures, and determine the costs associated with the new procedure.

2. __BQC Data__ - Since the beginning of the Quality Control/Random Audit program, benefit year earnings errors have contributed significantly to the QC error rate in our state. QC data continues to reflect this problem. During 1990, the largest number of both weeks and dollars overpaid were due to benefit year earnings errors. Specifically, 43% of weeks mispaid and 33% of dollars mispaid during 1990 were caused by BYE errors.

Analysis of specific benefit year earnings error cases indicates that the BYE problem is largely due to problems with the reporting of such earnings. Specifically, 50% of all BYE errors during 1990 were due to earnings being misreported. A case-by-case analysis of these errors indicates that nearly 80% were due to either claimants or employers misreporting earnings. The employer error resulted primarily from the failure to timely reply to the agency's request for earnings information. Claimants routinely lack documentation and improperly estimate their earnings.

While the QC error data clearly indicates a BYE reporting problem, we have done additional analysis which further substantiates the existence of a problem. As another measure of benefit year earnings reporting accuracy, we have analyzed QC data by identifying all cases in which any amount of earnings were reported before the QC investigation. We compared earnings before and after investigation, regardless of whether they reached the deductible threshold.

During 1990, 22% of the QC cases investigated had earnings reported before the investigation. Of these cases, 20% were determined to be incorrectly reported after the QC

174

investigation. So, even though our QC error rate data indicates that only 4% of weeks paid are in error due to BYE errors, we have evidence that 20% of weeks with earnings are reported incorrectly. Although many of these cases did not result in QC payment errors, the figures provide some indication of the extent to which earnings are not reported correctly.

3. Other UI Data - A substantial portion of the BYE reporting problem is due to the fact that employers fail to return reporting forms within the required time frames. A review done by our BPC Unit during 1988 indicated that approximately 50% of employers fail to return earnings verification forms within the required time frame. While it may be beneficial to address the reporting problem by working to increase timely reporting by employers, it is our view that the misreporting problem will best be addressed by establishing a front end method which prevents the inaccurate reporting of earnings when claimants sign for the benefits.

Additional UI data has been reviewed to determine the extent of the problem of misreported earnings within our UI system. Analysis of Benefit Payment Control and Research & Statistics data indicates that during 1990, 50% of all overpayments set-up were the result of incorrect reporting of benefit year earnings. This amounted to in excess of 43,000 BYE overpayment cases. In addition, there were over 44,000 cases where misreporting of BYEs resulted in extensive processing to make supplemental payments.

4. Problem Identification Summary - Benefit year earning mispayments in our state are primarily due to the misreporting of earnings. We feel this is a problem significant enough to justify a radical change in the current reporting process.

The current process requires that the Agency solicit earnings information from employers after claimants report earnings for a claim week. Employers have 5 days to respond. If they do not respond in that time, payment is made based on claimants' earnings estimate. This method has been shown to have several weak points that will be addressed by the test procedure.

First, the 5 day period complicates the process and puts undue pressure on the employer. The test procedure will simplify the process for employers because they will be able to complete the necessary form when preparing their payroll.

Second, the current method requires claimants to "certify" earnings, however, they frequently are only making a rough guess because they are unprepared to provide an accurate estimate. With the test procedure, claimants will no longer be required to estimate their earnings; they will only be required to provide the employer statement summarizing their earnings for the claim period.

B. STUDY OBJECTIVES/EXPECTATIONS

1. **Objective(s)** - The objectives of this study will be to compare the current benefit year earnings verification procedure to a test procedure and determine if earnings are reported correctly at a significantly higher rate with the test procedure.

2. **Expected Insights/Value to UI Management** - This study is expected to provide valuable information to management. The current BYE reporting methodology is subject to error because employers are asked to respond in a very short time and because claimants are required to certify earnings from previous weeks of employment based on "guesstimates" which frequently are incorrect.

Although many of the mispayments resulting from incorrect reporting of BYEs are corrected when employers return the necessary forms, the process of setting up overpayments and reissuing underpayments is operationally expensive. Following completion of the study, UI management will have sufficient information to compare the effectiveness of two BYE reporting methods. If the test procedure can be shown to significantly increase the proper reporting of earnings, then UI management will have the facts necessary to validate a procedural/policy change which will improve the overall payment of UI benefits and reduce operational costs.

C. STUDY METHODOLOGY

1. **Study Design** - Local offices in our state are divided into eight (8) geographical regions, with each being managed by separate regional staff. During this study we will utilize all local offices in two of these regions to compare the current BYE reporting procedure to an alternative procedure. The two procedures are as follows:

   a. The current (control) procedure as previously described requires the agency to mail a verification form to employers when claimants indicate that they had earnings during a claim week.

176

b. The alternative (test) procedure will require claimants to secure a verification form prepared by the employer, and present this when filing a partial claim.

To test the effectiveness of the alternative reporting procedure we will use a quasi-experimental comparison group design. Using this design we intend to institute the test procedure in Region A and compare the error rate from this region to the rate from the control region, Region B.

We understand that a comparison of this nature may result in validity threats if there are substantial differences between the claimant and employer populations in the two regions, or if there are regional differences in the manner which claims are processed, but through extensive analysis we have identified two regions that are similar in relation to relevant variables. We have determined that these regions are sufficiently similar to allow for our comparisons.

Rather than require all employers in Region A to modify their earnings reporting methods, we have identified the employers that employ the majority of claimants signing for partial benefits. We will target these employers to institute the test procedure. These employers account for 80% of partial payments in Region A and they will constitute the test group of employers. These employers will be familiarized with the partial benefits process, trained in the use of the necessary forms and asked to do the following:

a. Complete the forms for employees who indicate they will be signing for partial benefits;

b. Provide a payroll stuffer to employees to inform them of the new procedure for reporting UI earnings.

2. Sampling Plan - To compare the test and control procedures we will randomly sample from the population of partial payments made in the two regions during a 13 week study period. In a manner similar to the QC sampling method, we will create two sampling populations at the completion of each week. One of these groups will be comprised of all claimants from the test region that received partial benefits during the previous week (key week) from an employer utilizing the test reporting procedure. A unique adjustment code will be used to distinguish these partial payments from others, thus allowing us to create an uncontaminated target population. The other target population will consist of all claimants that claimed partial benefits in the control Region during the previous week.

These populations will be sorted by two key variables: WBA and amount paid, with WBA being the primary sort key and amount paid during the key week being the secondary key. From each population a random sample of 30 cases will be drawn each week for investigation.

We will sample fifteen cases per week from each group for 13 weeks. This will yield two samples of 390 cases. Sample sizes of 390 were selected based on a worst case estimate of 50% of cases being incorrect, a desired confidence level of 95% (z = 1.96), and a desired level of precision of +/-5%. This calculation was based on the following formula for determining sample size:

sqrt(n) = population variability * Z-score * 1/level of precision

Using this formula the sample sizes were selected as follows:

sqrt(n) = .5 * 1.96 * 1/.05
sqrt(n) = .5 * 1.96 * 20
sqrt(n) = 19.6
n = 384

Although the recommended sample size is 384, we have rounded that to 390 for each group so that we can draw weekly samples of 15 cases.

3. Study Variables

a. Dependent Variable - The dependent variable in this study is the proportion of partial payment cases in the test and control groups with payment errors. There are several levels of the dependent variable. For both the test and control groups they are:

(1) Proportion of weeks and dollars overpaid

(2) Proportion of weeks and dollars underpaid

(3) Total mispayments

During our audits we will track these levels of the dependent variable, and subsequent data analysis will compare the respective rates for the test and control regions.

b. Independent Variable - The test and control BYE reporting methods represent the independent variable. The test reporting procedure will be implemented in Region A, while the control region, Region B, will continue to process partial claims in the current manner.

c. Control Variables - During data collection we will record relevant data on SIC, employer size, claimant age and education, and local office claims load to use as control variables during the final analysis.

4. Data Analysis Plan - Data analysis will require use of a confidence interval for evaluating differences.

Because we are utilizing a quasi-experimental design that relies upon test and comparison groups, and probability sampling of a target population, our data analysis will be concerned mainly with inferential statistical analysis. In other words, the sample population findings will be extrapolated to the target population of all intrastate benefit year earnings cases.

We believe that the study's methodology adequately controls for relevant internal and external validity threats. We will construct 95% confidence intervals at the .05 level of precision for both the test and comparison group error rates. This will enable us to fairly evaluate the benefits of instituting the new procedure. If possible, we also anticipate construction of a linear model to assess the impact of various control variables and increase the precision of our findings.

5. Resource Estimation and Workplan - Attachments 1 through 4 are schedules which identify the major tasks and identify the expected time frames for their completion. In addition, Charts 1 and 3 identify available personnel resources and estimate the usage of these in relation to the major study tasks.

These charts establish that the study will require approximately 775 person days to complete. We expect this activity to begin on October 1, 1992 and conclude with our final report of findings on March 31, 1993. During this period, Quality Control investigators assigned to the study will continue to complete out-of-state requests and out-of-area requests from other investigators in a normal manner.

During the study period identified above, we are requesting the diversion of the equivalent of six (6) full-time Quality Control positions. We will utilize these positions in the following manner:

ACTIVITY	# OF QC POSITIONS
(1) Data analysis and report preparation	.5 QC Data Analyst position
(2) Case investigation, local office training, employer training	5 QC Investigator positions
(3) Data Entry and Project Support	.5 QC clerical position
TOTAL	6 QC positions

Given our above estimation of required resources, we are requesting quality control sample be reduced by 10 cases per week during the study period. This reduction will free five (5) Quality Control investigators and still maintain our quality control sampling at the required rate of 400 cases per year.

Your prompt approval of this PI study plan will be greatly appreciated. If there are any technical questions regarding this proposal, please contact our Quality Control data analyst at (717) 777-5252.

CHART 1
PERSONNEL RESOURCES AVAILABLE

ACRONYM	DESCRIPTION	AVAILABILITY (Person Days)
FCSPEC	Forms Control Specialist	1
SYSANAL	Systems Analyst	15
RESANAL	Research Analyst	72
DEOPR	Data Entry Operator	15
INVES	Investigator	650
LODEP	Local Office Deputy	22

CHART 2
MAJOR ACTIVITIES FOR THE BYE PI STUDY

ACRONYM	DESCRIPTION	DURATION	COST (PERSON DAYS)
DEVFORM	Design employer form	1 wk	5
TESTFORM	Field test employer form	1 wk	5
SELEMP	Select sample of employers	12 wks	15
DISTFORM	Distribute forms to employers	2 wks	10
DEVPROC	Develop written procedures for local offices	2 wks	10
TRNSTAFF	Train local office staff	2 wks	10
TESTPROC	Field test local office procedures	13 wks	5
COLDATA	Collect data in local offices	13 wks	10
INVESCLM	Investigate selected claims	13 wks	650
CODEDATA	Encode data	3 wks	15
ANALDATA	Analyze data	4 wks	20
FNLREPT	Prepare final report	4 wks	20

BLANK PAGE

CHART 3
ESTIMATED RESOURCE USAGE FOR MAJOR ACTIVITIES

JOB (ACTIVITY)	RESOURCE PERSON DAYS REQUIRED FOR ACTIVITY					
	FCSPEC Qty	SYSANAL Qty	RESANAL Qty	DEOPR Qty	INVES Qty	LODEP Qty
DEVFORM	1		4			
TESTFORM			5			
SELEMP		15				
DISTFORM			5		5	
DEVPROC			5			5
TRNSTAFF			5			5
TESTPROC			3			2
COLDATA						10
INVESCLM					80	
CODEDATA			5	15		
ANALDATA			20			
FNLREPT			20			

CHART 4
GANTT CHART FOR MAJOR ACTIVITIES FOR EXAMPLE PI STUDY

JOB (ACTIVITY)	DATE SCHEDULED*																		
	OCTOBER					NOVEMBER					DECEMBER				JANUARY				
		5	12	19	26	2	9	16	23	30	6	13	20	27	4	11	18	25	
DEVFORM		[-]																	
TESTFORM				[---	---]														
SELECTEMP		[---	---	---	---	---	---	---	---	---	---	---	---]						
DISTFORM														[-]					
DEVPROC		[---	---]																
TRNSTAFF				[---	---]														
TESTPROC						[-]													
INVESCLM							[---	---	---	---	---	---	---	---	---	---	---	---	--->
CODFDATA																[---	---	---	--->
ANALDATA																			-->
FNLREPT																			-->

CHART 4 - PAGE 2
GANTT CHART FOR MAJOR ACTIVITIES

<table>
<tr><th rowspan="3">JOB
(ACTIVITY)</th><th colspan="19">DATE SCHEDULED*</th></tr>
<tr><th colspan="5">FEBRUARY</th><th colspan="5">MARCH</th><th colspan="4"></th><th colspan="4"></th><th></th></tr>
<tr><th></th><th>5</th><th>12</th><th>19</th><th>26</th><th>2</th><th>9</th><th>16</th><th>23</th><th>30</th><th>6</th><th>13</th><th>20</th><th>27</th><th>4</th><th>11</th><th>18</th><th>25</th><th></th></tr>
<tr><td>DEVFORM</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>TESTFORM</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>SELECTEMP</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>DISTFORM</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>DEVPROC</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>TRNSTAFF</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>TESTPROC</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>INVESCLM</td><td></td><td>--]</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>CODEDATA</td><td></td><td colspan="2">[-----]</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>ANALDATA</td><td></td><td></td><td></td><td colspan="3">[----------]</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>FNLREPT</td><td></td><td></td><td></td><td></td><td></td><td></td><td colspan="3">[----------]</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

BLANK PAGE

NAME
Regional Administrator
U.S. Department of Labor
Employment and Training Administration
STREET
CITY, STATE ZIP

Dear :

We have been working in conjunction with members of your staff to develop a Quality Control Program Improvement Study. I am pleased to inform you that we have completed a proposal for a study of the effectiveness of our Benefit Rights Interview. This proposal (enclosed) is being submitted for your review and submission to the National Office.

I feel this study will provide valuable information from which we will be able to make an informed decision regarding the effectiveness of our current procedure. If the study reveals that a procedural change is warranted, it would be a major step forward in improving our UI system and service to our unemployed workers.

Please feel free to contact me at your convenience, if additional information regarding this study is necessary.

Sincerely,

SESA Administrator

Enclosure

(State) Department of Labor & Industry
Bureau of Unemployment Compensation
Quality Control Section

ABBREVIATED SAMPLE PROGRAM IMPROVEMENT STUDY PROPOSAL

BENEFIT RIGHTS INTERVIEW (BRI):

Adequacy Assessment of Current Procedure

Reference - UIPL 37-91 (August 13, 1991)
- Technical Assistance Guide

Explanation: This draft PI study proposal is an abbreviated form of a real proposal. It consists of dot pointed basic topices that should be addressed when preparing a PI proposal. A formal proposal must be in full narrative form.

I. PROBLEM IDENTIFICATION

A. Quality Control Data

* Our current method of providing Benefit Rights Information (BRI) is to provide claimants with a "rights and responsibilities" pamphlet.

* Claimants are required to read the pamphlet and certify in writing that they have read and understand its contents.

* Over time, Quality Control data has provided some evidence that the BRI is not effectively imparting required knowledge to claimants.

* Multiple Quality Control keyweek error cases contained statements by claimants that they were unsure of their responsibilities regarding the proper reporting of earnings, availability issues and separations from part-time employment.

* All Quality Control keyweek error cases folders for CLY 1991 were reviewed to identify supporting evidence of this problem.

* All CLY 1991 Quality Control cases indicating previously adjudicated nonmonetary issues were reviewed for supporting evidence of this problem.

* Supporting data was accumulated from case summaries and claimant nonmonetary determination statements contained in case folders.

* A total of 45 such cases were intensively analyzed for possible indication of substandard claimant UI knowledge.

* Five cases were found to have specific reference to the claimant's misunderstanding of his/her responsibilities.

* Two cases clearly indicated the claimant's misunderstanding of the proper reporting of Benefit Year Earnings.

* One case involved the claimant misunderstanding of "what" constituted a base period.

*   In one case the claimant indicated a lack of knowledge regarding the requirements for the proper submission of mail claims.

*   In the final case the claimant disclosed that he/she is functionally illiterate.

B. Other Data Sources Consulted

*   Samples of nonmonetary determinations are routinely reviewed and graded utilizing the Quality Performance Index by the Adjudication Section within our agency.

*   This section was consulted regarding claimant statements indicating a lack of UI knowledge.

*   This unit provided qualitative input that lack of understanding of UI responsibilities is "often" cited by claimants in nonmonetary determinations, and functionally illiterate claimants are occasionally identified during their reviews.

*   The Division of Internal Audits performs local office administrative reviews which include assessing the quality of selected nonmonetary determinations. This Division's local office audit close-out documents disclose information similar to that received from the Adjudication Section.

*   The Benefit Payment Control Unit reviewed a sample of 100 overpayments established during CLY 1991 and found that claimants cited a misunderstanding of their responsibilities as the primary reason for the overpayment 20% of the time.

C. Conclusion

*   Based upon Quality Control case-by-case review, the potential exists that an ineffective BRI is significantly contributing to the incidents of errors within the UI payment process.

*   Other data sources provide qualitative supporting evidence for this conclusion.

*   Presentation of these findings to our corrective action committee and UI Director was met with some skepticism that the current BRI has been clearly shown to be ineffective.

*   However, all agreed that additional data should be collected regarding the effectiveness of the pamphlet BRI method.

* The Quality Control Division, in conjunction with central office technical units, was asked to prepare a PI study proposal.

II. STUDY OBJECTIVES

* A general meeting was held with the UI Director, Chief of Benefits, Claims Program managers and selected local office staff to explain the problem and develop feasible study objectives.

* It was agreed that the primary research question to be answered by the study will be "how effective is the current BRI in informing new claimants of their UI rights and responsibilities?"

* This question implies either measuring the level of knowledge imparted by the BRI or assessing the accuracy of benefit claims filed subsequent to the claimants' certification of understanding.

* All parties agreed that attempting to draw a causal relationship between the accuracy of weekly claims and the BRI was impractical.

* The longitudinal nature of such an undertaking would be far too extensive and would present a number of uncontrollable validity threats.

* All parties agreed that measuring the claimants' level of knowledge was a more feasible approach to assessing the effectiveness of the BRI method.

* This approach will minimize study resources and provide a reliable assessment of the relationship between the BRI and claimant UI knowledge.

* Our specific objective will be to test claimant knowledge at the time of first payment certification.

* The study will provide specific, quantifiable data regarding the effectiveness of the pamphlet BRI method of providing claimants with adequate knowledge of their UI responsibilities.

III. METHODOLOGY

* A descriptive, single group time series design was selected to accomplish our objective.

* This type of methodology is best suited to evaluation of the status quo.

* A survey research test will be developed to provide specific information on the level of claimant UI knowledge.

* The approach will be operationalized in 40 randomly selected local offices statewide.

* Only new claimants that have not filed for UI benefits in the previous calendar year will be included in the study.

* This group of claimants will form our weekly population of interest.

* We recognize selection as a possible validity threat in this approach, but feel it is not significant.

* The study approach should control for the impact of history, maturation and statistical regression on the study results.

* We feel most external validity threats will be controlled by randomization of location.

A. Sampling Plan

* Claimants will be randomly selected from the target population on a weekly basis.

* Historically, the average weekly population of first payments amounts to 3,500 in the 40 randomly selected offices where the study will be conducted.

* Review of the UI claims data base indicates that 1,500 of these claims meet the population criteria on a weekly basis. This amounts to a 26 week population of 39,000.

* Utilizing the sampling table found in Appendix C of the Technical Assistance Guide, it was determined an appropriate sample size would be 381 cases.

* This sample size will provide a confidence level of 95% at a precision level of +/- 5%.

* The sample size will be administratively increased to 390 to allow for a static 26 week sample of 15.

* Claimants that meet the population criteria and file for their waiting week and/or first compensable week will be

192

arrayed in ascending order, first by WBA and then by social security number.

* We will utilize the CLY 1990 Quality Control table of random start numbers, calculate the slip interval and select a probability sample of 15 cases each week for 26 weeks (total n = 390).

B. Dependent Variable

* The dependent (outcome) variable in this study is the claimants' knowledge of UI responsibilities.

C. Independent (Cause) Variable

* The primary independent variable is the BRI method which requires claimants to read a rights and responsibilities pamphlet and certify their comprehension of the materials.

D. Control Variables

* The following data items will be collected as control variables during the study: claimant age, sex, education level, number of previous claims, number of previous nonmonetary issues and claimant occupation.

IV. DATA ANALYSIS PLAN

* The data analysis plan for this study will primarily involve a descriptive analysis of the dependent variable.

* Frequency distributions will be prepared and analyzed for the total sample population and the sample population falling within each control group.

* Distributions will be summarized using standard measures of central tendency.

V. DATA COLLECTION

* Cases selected will be assigned at a rate of 2.5 cases per week per investigator.

* Investigators will personally contact each claimant to schedule a claimant interview.

* Claimants will be asked to verify all control information.

*   Claimants will be required to answer a comprehensive study questionnaire without investigator assistance.

*   The study questionnaire will ask specific close-ended questions regarding the claimant's understanding of his/her UI responsibilities.

*   The answers to this questionnaire will be codified to form the basis for measuring the effectiveness of the current BRI method.

VI. EXPECTED RESULTS

*   The sample data is expected to provide valid measurements of the effectiveness of the current BRI method.

*   These findings will be presented in a comprehensive written final report that will provide the UI Director with sufficient information to determine whether or not the current BRI is meeting claimant and agency needs.

*   It is anticipated that if the study indicates the current BRI method is ineffective that alternative BRI methods will be explored.

VII. STUDY RESOURCES/WORKPLAN

*   The attached primary task worksheets identify major study activities and estimate staff required to complete them.

*   The attached Gantt Chart identifies the time-frames established for completing the study tasks which will take 12 months to complete.

*   The data collection phase of the study will require use of six (6) full-time Quality Control positions for a period of 9 months.

*   The final three (3) months of the study will be the analysis stage, and require the use of three (3) full-time Quality Control positions.

*   The analysis stage will consist of final data collection, data coding, data analysis and final report preparation.

*   Use of other SESA resources will be minor and limited to mainframe processing units for population maintenance and sample selection, local office staff for records retrieval and claimant orientation regarding the project.

- We will begin the study October 1, 1992 and complete the effort with submission of our final report on September 30, 1993.

APPENDIX C

CONTENTS

Page

Design 1: Pretest-Posttest Control Group Design. C-1

Design 2: Posttest Only Control Group Design. C-3

Design 3: Non-Equivalent Control Group Design C-5

Design 4: Single Group Time Series Design C-7

Design 5: Time Series with Non-Equivalent Control Group Design. C-9

Design 6: One Group Pre-test-Posttest Design. C-10

Flowchart of Designs for Measuring Cause/Effect Relationship. C-12

Possible Sample Sizes for Estimating Proportions C-13

196

DESIGN 1

NAME: Pretest-Posttest Control Group Design

TYPE: Experimental

DESIGN DIAGRAM

R O X O (treatment group)
R O O (control group)

R = randomly assigned subjects
O = observation (pretest/posttest)
X = treatment (independent variable)

DESIGN DESCRIPTION

This is the classic experimental design. Properly used it allows the researcher to control exposure to the independent variable, determine statistical association between variables, and control for other possible causal factors. Properly utilized, this design provides the most reliable evidence of a causal relationship.

The design requires that members of the study population be randomly (R) assigned to study and control groups. A pretest measure is taken for each group (O) and the study group is subsequently exposed to the treatment (X). Posttest measures (O) are then taken to assess the impact of the treatment.

BRIEF EXAMPLE

A random sample is selected from the population of all claimants that open a new claim during a given week. All sampled claimants are given a pretest to assess their knowledge of the UI rights and responsibilities. Following the pretest, the sampled claimants are randomly placed into test and control groups. The test group is given a video BRI, while the control group participates in a BRI in which they are provided a pamphlet explaining rights and responsibilities, asked to read the pamphlet, and given an opportunity to ask questions. After the interviews, a posttest is conducted to assess each groups' the understanding of each group in regard to their UI rights and responsibilities. Comparisons are made between the test and control groups to determine if one of the methods is a more effective method of informing claimants of UI rights and responsibilities.

VALIDITY CONSIDERATIONS

If carried out properly, this design is internally valid. External validity may be compromised, however, if there is an interaction between the pretest and the treatment which produces effects that do not develop within the control group.

DESIGN 2

NAME: Posttest Only Control Group Design

TYPE: Experimental

DESIGN DIAGRAM

R X O (treatment group)
R O (control group)

R = randomly assigned subjects
O = observation (posttest)
X = treatment (independent variable)

DESIGN DESCRIPTION

The posttest only control group design involves randomly assigning subjects to test and control groups, exposing the test group to a treatment, and taking posttest measures for both groups. This design is appropriate when the nature of the study does not require a pretest, or when one wants to eliminate the possibility that the pretest will in some way bias study results.

BRIEF EXAMPLE

A SESA recognizes that it has a problem with the manner in which new employers report base period wages. They plan to begin sending information packets to new employers with the expectation that the additional information will help new employers correctly complete BP wage reports. Before adopting this as a new procedure, however, the SESA would like to determine the effectiveness of the information packet.

The SESA maintains a record of new employers as they register with the Agency. At the end of each week, the population of new employers that started business that week is sorted by SIC code. From this sorted file random samples of new employers are selected to be in either the test or control groups. A skip interval is calculated and a systematic random sample is selected with the employer sampled by the skip interval being placed in the treatment group (scheduled to receive the information packet) and the employer immediately following (skip interval plus 1) being placed in the control group which will not receive the information packet.

After the wages are reported for a quarter, a random sample of employees from the treatment and control employer groups is sampled and audits completed. Measures are taken to determine the accuracy of the BP wage report in relation to the reporting of wages and credit weeks.

After each employer audit the investigator scores the employer as either correctly or incorrectly reporting BP wage information (this is the posttest).

VALIDITY CONSIDERATIONS

Internal Validity - Most threats to internal validity are controlled by the design. The presence of a randomly assigned control group eliminates history, maturation and selection as threats to internal validity. Testing, statistical regression, and instrumentation are eliminated as internal validity threats because there is no pretest.

External Validity - There is a possible selection problem if results from the study are to be generalized to the population of all employers. Because there may be systematic difference between new and existing employers, it is possible that the results of the study may not be generalizable to all employers.

DESIGN 3

NAME: Non-Equivalent Control Group Design

TYPE: Quasi-Experimental

DESIGN DIAGRAM

O X O
O   O

O = Observation (pretest/posttest)
X = Treatment (independent variable)

DESIGN DESCRIPTION

Although this design may look the classic experimental design, it differs because subjects are not randomly selected for exposure to the treatment. Rather, a treatment is presented to a test group, and a comparison group of similar subjects that has not been exposed to the treatment is selected and used for comparative purposes. This is a useful design for those situations where the nature of the evaluation precludes random assignment of individuals to a treatment, program or policy.

BRIEF EXAMPLE

A SESA's UI system is divided into 5 regions. Each region is managed by its own group of regional managers. One of the regions (Region 1) has implemented a new method of reporting benefit year earnings. Whereas the standard state procedure is to verify earnings via a mail verification to employers, Region 1 has modified the procedure so that claimants are required to submit certified earnings statements from the employer at the time of signing for partial benefits. The SESA is interested in evaluating the effectiveness of the two procedures.

Because the program has already been implemented in Region 1, the decision is made to compare Region 1 with a similar region. Region 2 is determined to be similar to Region 1, and is selected to serve as the comparison region. As part of the process of gauging the similarity of the two regions, figures are generated which indicate the benefit year earnings error rate for the two regions. These error rates constitute the pretest measure. Random samples are taken from each region and audits of partial payments are conducted to determine the error rates in each region. These error rates constitute the posttest measure.

Because random samples are taken from each region it would be easy to think that this is really an example of a classic experimental design. This is not the case, however, although test and control groups are randomly sampled, they are sampled from different regions. Consequently, claimants from Region 2 (comparison region) have no possibility of being included in the test group. Because claimants are not randomly exposed to the treatment, this is not a classic experimental design.

VALIDITY CONSIDERATIONS

The major limitation of this design is its inability to control for biases introduced by selection. Because the comparison group is selected based on assumed similarities between groups, it is possible that unknown systematic differences between groups may adversely affect results.

DESIGN 4

NAME: Single Group Time Series Design

TYPE: Quasi-Experimental

DESIGN DIAGRAM

O O O O O X O O O O O

O = Observation (pretests/posttests)
X = Treatment (independent variable)

DESIGN DESCRIPTION

The single group time series design is characterized by a series of observations at periodic intervals preceding introduction of the independent variable. After the independent variable (policy, program, procedure) is introduced, observations are continued periodically to assess both the effect of the treatment and the duration of the treatment effect.

In these designs the independent variable may be introduced by the researcher, but in many cases it is something which occurs naturally or is introduced by someone else. This design works best if the independent variable is expected to have a marked and immediate effect.

BRIEF EXAMPLE

A state introduces a new method of selecting QC samples. Rather than drawing statewide QC samples, the state now draws samples which are stratified by region (stratification is the independent variable). In an attempt to evaluate whether the revised sampling methodology has reduced the time required to complete QC cases, the state calculates quarterly averages for the time required to complete a case for the 6 quarters prior to the sampling modification (these are the pretest measures).

After introducing the sampling modification, the state continues to calculate quarterly averages for case completion time to see if it now takes less time to complete a case (these constitute the posttest measures).

DATA ANALYSIS

Data from single group time series designs can be analyzed by calculating summary statistics (mean, error rate, etc.) and graphing the data.

VALIDITY CONSIDERATIONS

The internal validity of this design is particularly vulnerable to the effects of history because no comparisons are made to a control group. If any other environmental changes occur that might produce the observed results, the analyst is unable to sort out the effects of the competing explanations.

Because there is no control group, this design does not control for external validity threats.

DESIGN 5

NAME: Time Series with Nonequivalent Control Group

TYPE: Quasi-Experimental

DESIGN DIAGRAM

```
O O O O O X O O O O O
O O O O O   O O O O O
```

O = Observation (pretests/posttests)
X = Treatment (independent variable)

DESIGN DESCRIPTION

This design differs from the single group time series design in that a comparison group has been added to control for the effects of history as an internal validity threat. This design is very much like an expanded non-equivalent control group design (Design 3).

BLANK PAGE

DESIGN 6

NAME: One Group Pretest-Posttest Design

TYPE: Nonexperimental

DESIGN DIAGRAM

O X O (test group)

O = observation (pretest/posttest)
X = treatment (independent variable)

DESIGN DESCRIPTION

This nonexperimental design uses pretest and posttest scores from a single group in an attempt to determine the effect of the independent variable. Because there are no control or comparison groups, however, the design is ineffective as a means of establishing causation because there is no way of knowing what results may have been obtained without the intervention.

BRIEF EXAMPLE

Suspecting that local office personnel were incorrectly coding countable nonmonetary determinations as noncountable, a SESA conducts a study which attempts to address the problem. The SESA selects a random sample of determinations and calculates the percentage of countable determinations which are erroneously coded as noncountable (this is the pretest measure). The SESA then institutes a training program to improve local office adjudicators' understanding of the criteria for countable and noncountable determinations (this is the independent variable). The SESA then selects another random sample of determinations to obtain posttest measures to evaluate the effectiveness of the training program.

VALIDITY CONSIDERATIONS

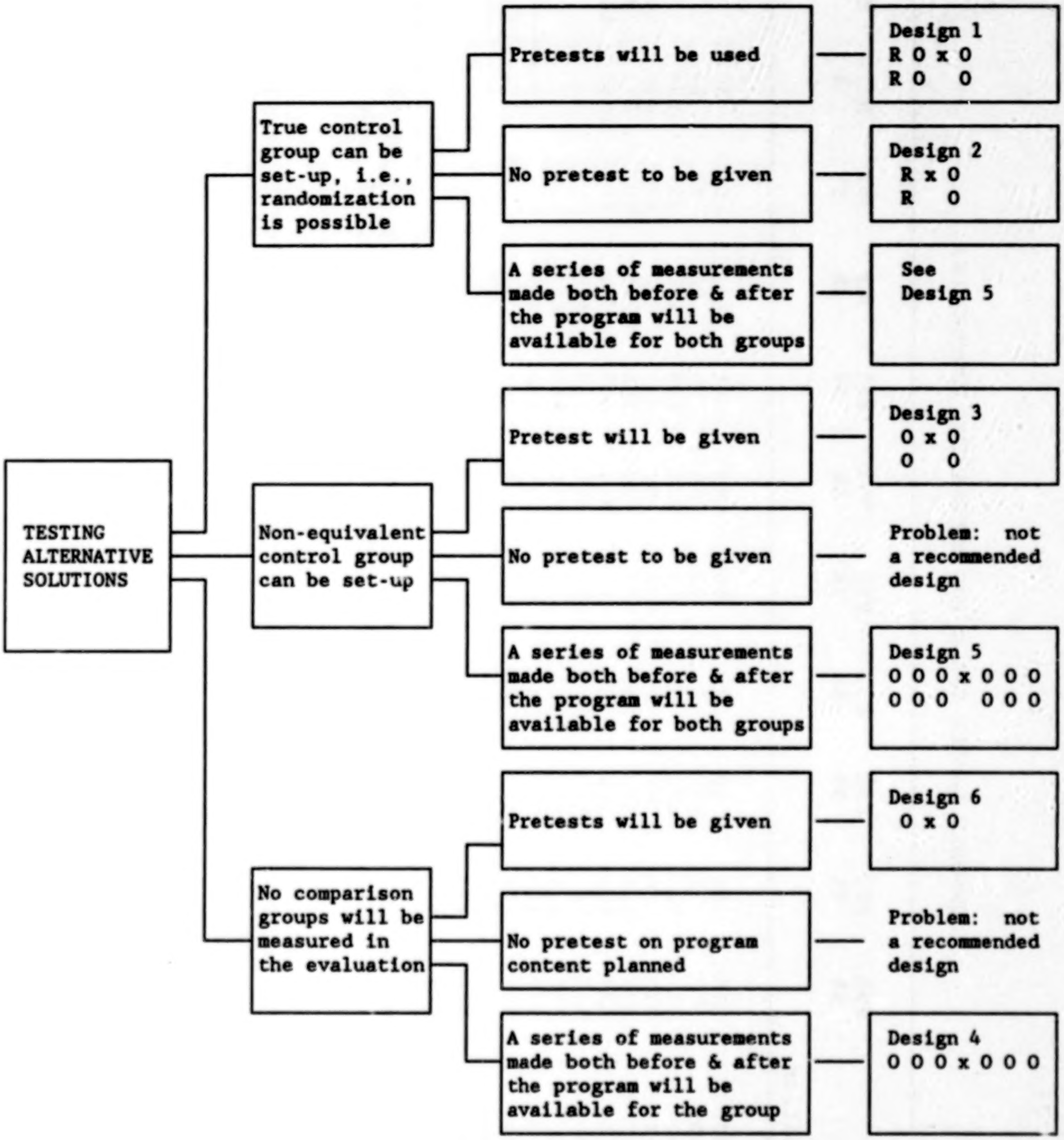
There are several weaknesses inherent in the design which may affect the internal validity of the findings. Specifically, since comparisons have not been made to a control group that did

207

not receive training, we cannot discount the possibility that some factors (increased experience, policy change, etc.) have not operated separately from the training to improve the accuracy with which adjudicators code their determinations.

*Design case studies adapted from O'Sullivan and Rassel, 1989. pp. 52-69.

FLOWCHART OF DESIGNS FOR MEASURING CAUSE/EFFECT RELATIONSHIPS


TESTING ALTERNATIVE SOLUTIONS
True control group can be set-up, i.e., randomization is possible
Pretests will be used
Design 1
R O x O
R O O
No pretest to be given
Design 2
R x O
R O
A series of measurements made both before & after the program will be available for both groups
See Design 5
Non-equivalent control group can be set-up
Pretest will be given
Design 3
O x O
O O
No pretest to be given
Problem: not a recommended design
A series of measurements made both before & after the program will be available for both groups
Design 5
O O O x O O O
O O O O O O
No comparison groups will be measured in the evaluation
Pretests will be given
Design 6
O x O
No pretest on program content planned
Problem: not a recommended design
A series of measurements made both before & after the program will be available for the group
Design 4
O O O x O O O

NOTE: This diagram assumes the evaluation is being planned before a program (treatment) has been implemented. For more information on evaluations planned after program implementation see. . .Source.

SOURCE: How to Design a Program Evaluation, Fitz-Gibbon & Morris, Sage Publications

POSSIBLE SAMPLE SIZES FOR ESTIMATING PROPORTIONS

CONFIDENCE LEVEL:	95%	90%	85%	80%	95%	90%	85%	80%	95%	90%	85%	80%
PRECISION:	5%	5%	5%	5%	3%	3%	3%	3%	1%	1%	1%	1%
Size of Population												
50	44	42	40	39	48	47	46	45	50	50	50	49
100	79	73	167	63	92	88	85	82	99	99	98	98
200	132	116	102	92	169	159	148	140	196	194	193	191
500	217	178	147	126	343	302	268	242	476	466	456	447
1,000	278	216	172	145	521	434	365	319	907	873	838	809
2,000	322	242	188	156	705	554	447	380	1661	1550	1443	1357
5,000	357	261	199	163	894	664	516	429				
10,000	370	268	203	166	982	711	545	448				
20,000	377	272	205	168	1033	737	560	459				
50,000	381	274	207	168	1066	754	569	465		2,000		
100,000	383	275	207	169	1077	760	573	467				
1,000,000	384	275	207	169	1088	765	576	469				
100,000,000	384	275	207	169	1089	765	576	469				

Example of use of table: If you are sampling from a universe of 10,000 and you want to be 95% confident that your results will be within 5% of the true percentage in the population, you need a random sample of 370.

*Note: Table adapted from various sources of such tables.
**Note: A random sample size of 2,000 is generally considered sufficient for 1% precision regardless of population size.
***Note: The creation of this table assumes a worst case population variance of 50% (.5).

☆ U.S. GOVERNMENT PRINTING OFFICE: 1992—312-414/64848

END

DATE FILMED

07 / 31 / 95